| | |
|---|---|
| Title | Operant Variability:An Experimental Analysis of Consequential Control over Response Variation |
| Author(s) | Murray, Clodagh Mary |
| Publication Date | 2012-09-29 |
| Item record | http://hdl.handle.net/10379/3286 |

# Operant Variability: An experimental analysis of consequential control over response variation.

Clodagh Murray

Supervisor: Dr Olive Healy

School of Psychology

National University of Ireland, Galway

September 2012

Submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy

**Summary**

This thesis investigated the differences in response variability between children and adolescents with a diagnosis of autism spectrum disorder (ASD) and neurotypical children and adolescents. It also examined methods for increasing variability in children and adolescents with ASD using automated programs and, subsequently, naturalistic toy play settings.

Chapter 2 found that there were significant differences between the groups on a naturalistic test of variability. Furthermore, low variability on this test was found to be inversely related to overall rates of restricted repetitive behaviours (RRBs) and, in particular, the class of RRBs referred to as Repetitive Sensory and Motor Behaviours. Chapter 3 confirmed this difference in variability, this time using an automated procedure which utilised the U-value statistic as a measure of uncertainty, or variation on the data. A significant negative relationship between U-value and age was revealed, contrary to previous findings that reported a positive relationship.

Chapter 4 incorporated a sophisticated computer program which the participants interacted with as a game. Some six conditions were presented according to different lag schedules and the effect of the changing lag requirement on levels of variability in responding was measured. Significant differences in variability were observed between the control condition and the lag 6 and lag 8 conditions, indicating that the greatest variability is obtained in the presence of the higher lag values. However, the lag 8 condition produced slightly less variability than the lag 6 condition, suggesting that there may be a ceiling effect when using lag schedules with human participants.

Chapter 5 contained two experiments where the procedures investigated in the previous chapters were used to increase variability in a socially significant behaviour, play actions with toys. Significant improvements were noted for all five participants with ASD and suggestions are made for incorporating such procedures into intervention programmes for this population.

**Table of Contents**

Declaration…….vii

Acknowledgements…….vii

List of tables and figures……..ix

Chapter 1: Operant Variability Literature Review

1.1. Variability as operant behaviour……2

1.1.1. Sensitivity to contingencies of reinforcement…….2

1.1.2. Control by discriminative stimuli……4

1.2. Why is variability important?.....7

1.2.1. Importance of variability for problem solving……8

1.2.2. Importance of variability in the development of play and social behaviour……...9

1.2.3. Importance of variability in intervention for ASD……10

1.3. Extinction induced variability……...11

1.4 Lag schedules of reinforcement…….16

1.5 Measures of variability……...18

1.5.1. U-value statistic……..18

1.5.2. Percentage of trials in which variability contingencies are met……....20

1.5.3. Other measures……20

1.6. Implications for learning and performance……20

1.6.1. Creativity…….20

1.6.2. Problem-solving……22

1.6.3. Verbal behaviour……..25

1.6.4. Play……...28

1.7. Generalisation of variability to untargeted repertoires……32

1.8. Evidence of response variability in ASD……34

1.9. Restricted repetitive behaviours and low variability in ASD…….38

1.10. Play and variability in ASD……..40

1.11. Summary……43

Chapter 2. Study 1: Response variability in children with autism spectrum disorder: comparisons with neurotypical children and relationships to restricted repetitive behaviour subtypes

2.1. Introduction……46

2.2. Method……48

2.3. Results…….54

2.4. Discussion……57

Chapter 3. Study 2: Investigating the differences in variability between children with autism spectrum disorder and neurotypical children: relationships to age and receptive language ability

3.1. Introduction……61

3.2. Method……65

3.3. Results…….75

3.4. Discussion……78

Chapter 4. Study 3: Increasing variability in children with autism spectrum disorder and neurotypical children using lag schedules of reinforcement

4.1. Introduction……84

4.2 Method……86

4.3. Results and discussion……...102

4.4. General discussion…….128

Chapter 5. Study 4, Experiment 1: Increasing variability in actions with toys for five children with a diagnosis of autism spectrum disorder using lag schedules of reinforcement

5.1. Introduction……..132

5.2. Method…….137

5.3. Results and discussion……..147

5.4. General discussion……...…….176

Chapter 5. Study 4, Experiment 2: Increasing variability in actions with a child diagnosed with autism spectrum disorder using incremental lag schedules of reinforcement

5.5 Introduction……..…….183

5.6. Method…….……184

5.7. Results and discussion……..190

5.8. General discussion……198

Chapter 6: General discussion: Operant variability: An experimental analysis of consequential control over response variation

6.1. Overview of aims…….202

6.2. Summary of thesis findings…….204

6.3. Theoretical implications of the current findings…….214

6.3.1. Implications for autism intervention…….216

6.5. Implications for future research……..219

6.6. Limitations of the current thesis…….221

6.6.1. Sampling and generality of findings……221

6.6.2. Standardised measures used…….222

6.6.3. Maintenance and follow-up……..223

6.7. Conclusion ……..223

References……...225

Appendix A……

**Declaration**

This thesis is the result of my own investigations. This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed…….. (candidate)

Date…….

**Acknowledgements**

First and foremost I would like to sincerely thank my amazing supervisor, Dr Olive Healy. Your passion for the field of Applied Behaviour Analysis has been a constant source of inspiration to me. Through your unsurpassed knowledge and expertise I have learned more than I ever could have hoped. I cannot thank you enough for all the support, encouragement and guidance you have shown me over the last four years.

To Joseph Mee and Yvonne Hennessey who brought my ideas to life with their fantastic programming and design skills, thank you for your time and effort, your work is hugely appreciated. Huge thanks also to Declan Coogan for his remarkable patience and his assistance with all things technological.

I would like to thank all of my friends for their continued support, in particular Michelle Kelly, friend and colleague. We started and finished this journey together and it would have been a much tougher road without you. Thank you for everything. To Jason, Jenny, Elizabeth and Lisa, your kindness is much appreciated.

To my family, my parents Tom and Fionnuala, sisters Maeve and Mairéad and my brother Michael, thank you for everything you have done for me. I could not have finished this thesis without your constant and unwavering love and support. You all mean the world to me and I know how lucky I am to have such a wonderful family.

Finally, to all of my participants and their parents and teachers, thank you for helping me with my research, without you none of this would have been possible.

**List of Tables and Figures**


| | |
|---|---|
| Page 53 | *Table 2.1:* Variability patterns and associated codes as outcomes of the Penny-Hiding Game. |
| Page 57 | *Table 2.2:* Kendall's tau correlations between variability and scores on the RBS-R. |
| Page 67 | *Table 3.1:* List of matched triads along with demographic information (age and sex). |
| Page 71 | *Table 3.2:* List of all 16 possible sequences of four presses using two buttons. |
| Page 76 | *Table 3.3:* Mean U-values and Standard Deviations for the three groups. |
| Page 88 | *Table 4.1:* Participant information for Group 1 (age, sex and ADOS-G scores) and Group 2 (age and sex). |
| Page 90 | *Table 4.2:* Description of each level of the computer program. |
| Page 93 | *Table 4.3:* List of all possible sequences of three presses across four keys. |
| Page 94 | *Table 4.4:* List of the four program settings with the order of conditions for each. |
| Page 102 | *Table 4.5:* U-values for the participants in Group 1 across all conditions, together with mean U-value for each condition. |
| Page 103 | *Table 4.6:* U-values for the participants in Group 2 across all conditions, together with the mean U-value for each condition. |
| Page 104 | *Table 4.7:* Descriptive statistics: Mean U-values and standard deviations for Group 1, Group 2 and Group 1/Group 2 combined for each of the six conditions. |
| Page 137 | *Table 5.1:* Participant characteristics |
| Page 138 | *Table 5.2:* List of training toys |

Page 139 *Table 5.3:* List of generalisation toys

Page 143 *Table 5.4:* Topographies of stereotypical behaviours for each participant.

Page 148 *Table 5.5:* ADOS scores and diagnoses

Page 148 *Table 5.6:* RBS-R scores by subscale

Page 174 *Table 5.7:* Number of prompts issued across training sessions for each participant.

Page 174 *Table 5.8:* Means and standard deviations for number of different actions in each phase.

Page 186 *Table 5.9:* Description of toys sets and their component toys used in each phase of the experiment.

Page 188 *Table 5.10:* Description of the order of conditions in each phase in the reversal design, where A = Baseline Sessions, B = lag Sessions (lag 1, 2 or 3, depending on phase) and gen = Generalisation Probes.

Page 196 *Table 5.11:* List of high-frequency actions and frequency at which they were observed in each session (when frequency was ≥5).

Page 55 *Figure 2.1:* Number of participants in Group 1 achieving each variability score (1-10).

Page 55 *Figure 2.2:* Number of participants in Group 2 achieving each variability score (1-10).

Page 72 *Figure 3.1:* Screen shot of computer program during trials.

Page 75 *Figure 3.2.* Mean age in months of EXP (experimental group), EXP AE (age equivalence scores of the EXP group), CCA (chronological age matched controls) and CAE (age equivalence matched controls).

Page 76 *Figure 3.3:* Mean U-value for each group (Experimental, Chronological Age controls and Age Equivalence controls).

Page 77 *Figure 3.4:* Mean standard score (SS) for the 3 groups (Experimental, Chronological age controls and Age equivalence controls.

Page 91 *Figure 4.1:* Apparatus: The laptop and Ergodex® with the program on display.

Page 96 *Figure 4.2:* One level of the program: bee travelling along the flowers to reach honey.

Page 96 *Figure 4.3:* One level of the program: mouse looking for cheese.

Page 97 *Figure 4.4:* One level of the program: rocket travelling across planets to reach a birthday party.

Page 105 *Figure 4.5:* Mean U-values for Group 1 across the six conditions.

Page 105 *Figure 4.6:* Mean U-values for Group 2 across the six conditions.

Page 108 *Figure 4.7:* U-value scores for Participant 1 across the six conditions.

Page 109 *Figure 4.8:* U-value scores for Participant 2 across the six conditions.

Page 110 *Figure 4.9:* U-value scores for Participant 3 across the six conditions.

Page 111 *Figure 4.10:* U-value scores for Participant 4 across the six conditions.

Page 111 *Figure 4.11:* U-value scores for Participant 5 across the six conditions.

Page 112 *Figure 4.12:* U-value scores for Participant 6 across the six conditions.

Page 113 *Figure 4.13:* U-value scores for Participant 7 across the six conditions.

Page 114 *Figure 4.14:* U-value scores for Participant 8 across the six conditions.

Page 115 *Figure 4.15:* U-value scores for Participant 9 across the six conditions.

Page 115 *Figure 4.16:* U-value scores for Participant 10 across the six conditions.

Page 118 *Figure 4.17:* U-value scores for Participant 11 across the six conditions.

Page 119 *Figure 4.18:* U-value scores for Participant 12 across the six conditions.

Page 120 *Figure 4.19:* U-value scores for Participant 13 across the six conditions.

Page 120 *Figure 4.20:* U-value scores for Participant 14 across the six conditions.

Page 121 *Figure 4.21:* U-value scores for Participant 15 across the six conditions.

Page 122 *Figure 4.22:* U-value scores for Participant 16 across the six conditions.

Page 123 *Figure 4.23:* U-value scores for Participant 17 across the six conditions.

Page 124 *Figure 4.24:* U-value scores for Participant 18 across the six conditions.

Page 125 *Figure 4.25:* U-value scores for Participant 19 across the six conditions.

Page 126 *Figure 4.26:* U-value scores for Participant 20 across the six conditions.

Page 127 *Figure 4.27:* U-values of four pigeons in REPEAT (open bars) and VARY (closed bars) under different lag schedules (reproduced with permission from Abreu-Rodrigues et al., 2005).

Page 139 *Figure 5.1*: Training toys

Page 139 *Figure 5.2*: Generalisation toys.

Page 150 *Figure 5.3:* Number of different actions demonstrated with toys by Participant 1.

Page 151 *Figure 5.4*: Number of different toys sampled in each session by Participant 1.

Page 152 *Figure 5.5*: Number of switches from toy to toy in each session by Participant 1.

Page 153 *Figure 5.6*: Frequency of stereotypical behaviours emitted per session by Participant 1.

Page 154 *Figure 5.7*: Results from baseline and post intervention probes with generalisation toys showing number of different actions, number of toys sampled, number of switches from toy to toy and frequency of stereotypy for Participant 1.

Page 155 *Figure 5.8*: Number of different actions with toys demonstrated by Participant 2.

Page 156 *Figure 5.9*: Number of toys sampled by Participant 2.

Page 157 *Figure 5.10:* Number of switches from toy to toy by Participant 2.

Page 157 *Figure 5.11*: Frequency of stereotypy emitted by Participant 2.

Page 158 *Figure 5.12*: Results from baseline and post-intervention probes with the generalisation toys showing number of different actions, number of toys sampled, number of switches from toy to toy and frequency of stereotypy for Participant 2.

Page 160 *Figure 5.13*: Number of different actions with toys demonstrated by Participant 3.

Page 161 *Figure 5.14:* Number of different toys sampled by Participant 3.

Page 162 *Figure 5.15*: Number of switches from toy to toy demonstrated by Participant 3.

Page 163 *Figure 5.16*: Frequency of stereotypy emitted by Participant 3.

Page 164 *Figure 5.17*: Results of baseline and post-intervention probe with generalisation toys for Participant 3.

Page 166 *Figure 5.18:* Number of different actions demonstrated by Participant 4.

Page 167 *Figure 5.19:* Number of different toys sampled by Participant 4.

Page 168 *Figure 5.20*: Number of switches from toy to toy by Participant 4.

Page 169 *Figure 5.21:* Results for pre and post intervention probes with generalisation toys for Participant 4.

Page 171 *Figure 5.22:* Number of different actions with toys demonstrated by Participant 5.

Page 171 *Figure 5.23:* Number of toys sampled by Participant 5.

Page 172 *Figure 5.24:* Number of switches from toy to toy by Participant 5.

Page 173 *Figure 5.25:* Results for pre and post intervention probes with generalisation toys for Participant 5.

Page 187 *Figure 5.26:* Toy sets used in Experiment 2 (clockwise from top left): Farm Set, Post Office Set, Pirate Ship Set and Ambulance Set.

Page 193 *Figure 5.27:* Number of different actions demonstrated with toys across all sessions (baseline, lag reinforcement and generalisation) for Phase 1, Phase 2 and Phase 3.

Page 194 *Figure 5.28:* Number of different actions demonstrated with toys in each of the generalisation sessions of Experiment 2.

# Chapter 1

# Operant Variability Literature Review

## 1.1.Variability as Operant Behaviour

It has been posited that variability is operant behaviour that is shaped and maintained by its consequences (Neuringer, 2002). By delivering reinforcement contingent on variable responding it is possible to increase the variability of an individual's behaviour, allowing them to react in more complex and varied ways to diverse situations. Machado (1989) reported that variability is a property of behaviour that shows sensitivity to its consequences and, therefore, is responsive to operant conditioning. Depending on the contingencies that are present in the environment, a high or low level of variable behaviour may be observed. Research from the experimental analysis of behaviour continues to illustrate that variability is an operant dimension and that it is possible to increase it through direct reinforcement (Lee, Sturmey & Fields, 2007). According to Miller and Neuringer (2000), typical behaviour manifests "differing levels of behavioural variability, which are controlled by discriminative stimuli and reinforcement contingencies" (p. 162), as is the case with the more commonly studied operant behaviours such as response rate, topography and force (Neuringer & Jenson, 2012).

**1.1.1. Sensitivity to contingencies of reinforcement.** There is much support for the hypothesis that variability is under the control of the consequences that follow it. Pryor, Haag and O'Reilly (1969) provided reinforcement to porpoises for jumping, swimming, turning or diving in novel ways. They delivered reinforcement when the porpoises emitted behaviour that they had not previously emitted. As a result of reinforcement of novel responses the porpoises began to emit a large range of varied and complex behaviours when they were in the experimental context and, after a while, they began to emit behaviours that had never before been observed in their species. Blough (1966) reinforced bar-pressing in pigeons when

they demonstrated the least frequent of a set of inter-response times. All three pigeons in this experiment were found to approximate a random distribution of inter-response times in their bar-pressing.

Miller and Neuringer (2000) determined that direct reinforcement of variable sequences of button-presses during a computer game could increase variability in individuals with a diagnosis of ASD and neurotypical controls. Initially, reinforcement was delivered independently of variable responding. Next, it was contingent on variability and then it was independent of variability again. All participants demonstrated significantly higher variability when reinforcement was provided contingently. Also, for all groups, variability remained high in the second phase when reinforcement again was delivered independently of response variation. This suggests a maintenance effect for reinforced variability in children with ASD and in neurotypical children and adults.

Balsam, Deich, Ohyama and Stokes (1998) awarded points to student participants for pressing buttons to navigate through a matrix on a computer. The task involved a 6 x 6 matrix with a white square in the upper left-hand corner, the start position, and a yellow square on the bottom right-hand corner, the finish position. The students were required to use “arrow down” and “arrow right” keys to move the square through the matrix from the white square to the yellow square. Points were awarded for sequences of five right and five down presses. If either key was pressed more than five times, the white box returned to the start position and no points were awarded. Four conditions were used: a “lag 0” condition where any sequence of five right and five down presses received points, a “lag 2” condition where the sequence had to differ from the previous two sequences in order to earn points, a “lag 10” condition where the sequence had to differ from the previous 10

sequences to access reinforcement and a “lag 25” condition where a sequence had to differ from the previous 25 sequences for points to be awarded. The results showed that as the variability criterion increased so did variability in responding on the matrix task. The authors concluded that: “It is thus possible to modulate variability by using variability as a criterion for reward” (Balsam, Deich, Ohyama & Stokes, 1998, p. 406). This provides further evidence for the operant nature of variability. It can be influenced by its consequences in the same way as other dimensions of behaviour.

Neuringer (2004) reported that reinforcement can “powerfully and precisely” control parameters of behaviour including which aspects of a response will be repeated, which aspects will vary and by how much. Ross and Neuringer (2002) illustrated this point effectively. They asked college students to draw rectangles on a computer screen and awarded points for variation in terms of size, form (height/width) and location on screen. They applied a different contingency to each of the three study groups so that for each group points would be contingent on simultaneously repeating along one dimension and varying along the other two e.g., maintaining the same location but varying size and form. The participants were not informed of the contingencies in place at any time but, nevertheless, the three groups differed significantly from each other, with each group responding appropriately to the contingencies in place for them. The precise influence of contingencies or schedules of reinforcement demonstrated by Ross and Neuringer (2002) helps to substantiate claims that variability should be considered as operant behaviour.

**1.1.2. Control by discriminative stimuli.** A number of studies have investigated the effects of using different discriminative stimuli (Sds) along with schedules of reinforcement to increase variability in animals, with much success

(Grunow & Neuringer, 2002; Neuringer, Deiss & Olsen, 2000; Page & Neuringer, 1985). Skinner defined the operant response as being controlled by discriminative stimuli as well as by consequences. Therefore, studies demonstrating that variability in responding can come under such control provide a significant step to further the acceptance of variability as operant behaviour among behaviour analysts and others.

Page and Neuringer (1985) delivered reinforcement to pigeons for repeating a single sequence (LRRLL) in the presence of a blue light and for emitting variable sequences in the presence of a red light. After every 10 reinforcements the red and blue lights were altered. The pigeons successfully learned to repeat sequences in the presence of the blue light and to vary sequences in the presence of the red light. The experimenters then reversed the contingencies, so that reinforcement was delivered for repeating sequences in the presence of the red light and varying sequences in the presence of the blue light, and the birds quickly reversed their patterns accordingly.

Further evidence for discriminative stimulus control over variability in responding was offered by Denney and Neuringer (1998). In their first experiment the authors delivered reinforcement to rats for emitting variable sequences of lever presses in the presence of one stimulus (the VAR condition). In the other condition (the YOKE condition) they delivered reinforcement contingent on pressing the levers but independently of variability. The frequency of reinforcement in the YOKE condition was matched to the frequency of reinforcement in the VAR condition. For half the rats “light off/tone on” signalled the VAR condition and “light on/tone off” signalled the YOKE condition. For the other half the respective discriminative stimuli were reversed so that “light off/tone on” signalled YOKE and “light on/tone off” signalled VAR.

An extra precaution was taken in this experiment to ensure that it was the discriminative stimulus that was exerting control over the behaviour. The sequences occurring immediately after the change of condition (and so the change in stimulus) were analysed separately from those occurring later on. If the external discriminative stimuli were influencing variability, differences would be seen immediately following stimulus onset when the only information available was provided by those external stimuli. This was in order to provide clearer evidence that it was the stimulus change that influenced responding and not some other factor. The results of this study showed that the discriminative stimuli exerted control over the rats' responding. When reinforcement depended on high variability in the presence of one stimulus, sequence variability was significantly higher than when reinforcement was independent of variability in the presence of a different stimulus. The greatest differences in variability were seen immediately following onset of the new discriminative stimulus.

In a second experiment, Denney and Neuringer (1998) designed three phases. During Sessions 1 and 3 the procedure was identical to that of Experiment 1 described above. In Session 2, the only difference was that the light and tone were continuously off. Thus, the contingencies of reinforcement were the same but the external discriminative stimuli were absent. Results showed that there was a significant difference in variability in the VAR and YOKE conditions of Sessions 1 and 3 but no difference in variability during the stimuli-absent session in between. The authors reported that in this experiment, if the first sequence did not result in reinforcement then variability increased in YOKE and decreased in VAR i.e., levels of variability converged in this session.

According to Denney and Neuringer (1998), control of variability by discriminative stimuli rules out another explanation of variability in behaviour – that which asserts that it is elicited by decreased reinforcement frequency or extinction (Stokes, 1995). While there is evidence that these can result in increased variability, such processes cannot control variable responding and maintain behaviour change in the specific way that direct reinforcement of variability can (Neuringer, 2002). This will be examined in more detail in the sections below.

### 1.2. Why is Variability Important?

Skinner (1938) reported that the process of operant shaping depends on a foundation of response variation to begin with. "Successive approximations to some goal response are selected for reinforcement and, without sufficient variation, selection is difficult or impossible" (Page & Neuringer, 1985, p. 429). Thus, the shaping process for all the skilled actions in the human repertoire relies on the ability to vary behaviour. Such variability provides the basis from which appropriate and adaptive behaviours can be shaped in the presence of the appropriate discriminative stimuli. Behaviour can be shaped only because there is always a distribution of responses to select from (Epstein, 1996). Variability may therefore play a significant role in the development of human behaviour.

Individuals frequently emit variable behaviour by combining two or more previously established behaviours in novel ways, or by combining old and new behaviours and techniques. These variable patterns of behaviour, i.e., connecting and combining concepts in new ways, provide evidence for the theory that higher variability of behaviour is adaptive for individuals and for society. Different repertoires of behaviour can come together in new situations to produce new

sequences or blends of behaviour or behaviours that have new functions**.** Variability can result in advances in science, education and the arts and it can also improve outcomes for the individual in terms of maximising potential and developing talents and skills.

According to Manoel and Connolly (1995) skills are created from the relationship between the organism, the environment and the task. As these three components are continuously changing, the skills required of an individual must be flexible enough to accommodate such changes, which are often subtle in nature. The authors also note that skilled actions are about the relationship between the means and the end, with the means being adjusted to ensure greater pay-off. Therefore, variability in behaviour is reinforced through goals being achieved and accessing such pay-offs. Without the ability to vary behaviour, individuals would not be able to attain their goals if features of themselves, their environment or their task were to alter even slightly. For example, if a person's mobility was decreased due to injury, they would adapt more effectively if they had the ability to vary their behaviour; if a teacher presented a previously mastered task in a different way, a child would require the ability to vary their behaviour in order to complete the task.

**1.2.1. Importance of variability for problem-solving.** The ability to behave variably is also an essential component of problem-solving. In order to solve problems, individuals must use their knowledge and skills in novel ways. Siegler (1995) investigated number conservation in five-year-old children. Some were given feedback on their performance, some were given feedback and asked to explain their performance and some were given feedback and asked to explain the experimenter's reasoning that led him to that feedback. The third group showed more learning than the other two groups. The data showed that the number of explanations proposed

during the pre-test phase significantly predicted percentage correct in the training sessions. Another significant factor was whether a child gave two or more explanations on a single item in the pre-test. Those with multiple explanations had higher percentage correct in the training phase than those who didn't. These two factors together accounted for 55% of the variance in percentage correct in the training phase. The author concluded that higher variability is positively correlated with problem-solving ability and skill acquisition in children.

**1.2.2. Importance of variability in the development of play and social behaviour.** The development of social behaviour in humans relies on the ability to emit variable responses. The importance of play in the development of social skills has been widely documented (Wenner, 2009). Children learn many important skills through play but in order for play to function in this way the child must have the ability to sample multiple materials and situations. They must also learn how to play with toys and peers by trial and error and feedback from adults and other children. If a child demonstrates low variability in their sampling behaviour and in their actions with toys they are likely to miss out on the opportunity to acquire basic repertoires of play as they cannot access the typical contingencies that continually shape play behaviours. In turn, this can mean that such children also miss out on the opportunity to learn early social skills that most children learn through play. For example, children with ASD have been shown to demonstrate low variability in their sampling behaviour (Mullins & Rincover, 1985) and in their actions with toys (Stahmer & Shcreibman, 1992). This may have important implications for their later social abilities.

Demonstrating variability in play is reinforcing for the individual and their peers (Neuringer, 2002) and it is likely that play serves as "variability training" for

later in life (Sutton-Smith, 1975). During play, children learn to respond to a variety of discriminative stimuli and receive reinforcement for doing so in novel and complex ways, thus laying the foundation for the development of social behaviour as children mature. The ability to vary behaviour in social situations is tremendously important. For example, individuals must vary their conversation in terms of content, tone and even volume according to the listener. Furthermore, the ability to differentiate between formal and informal social situations relies on the individual having a varied repertoire and also on their ability to use their social and conversational skills in varied ways according to ever-changing environments.

**1.2.3. Importance of variability in intervention for ASD.** Many authors have discussed autism spectrum disorder (ASD) in terms of "keystone deficits", which once remediated, would result in positive behaviour change across many areas of functioning (Mullins & Rincover, 1985; Rincover, Cook, Peoples & Packard, 1979). Mullins and Rincover (1985) suggested that teaching children to discriminate changes in reinforcement schedules may be one of the most important of these keystone deficits to target and they also recommend teaching children with ASD to effectively sample available options, thereby altering behaviour patterns before selecting the most appropriate one. This skill of responding variably would allow children with ASD to maximise the available reinforcement in a variety of environments.

Furthermore, a lack of behavioural variability can also be described as a keystone deficit within ASD. Teaching children with ASD to vary their behaviour across multiple repertoires may result in further behaviour change that is derived from this. Mullins and Rincover's finding that children with ASD are less likely to sample available options in their environment may be related to a lack of behavioural

variability in the area of exploration. Increasing variability in this population should result in children with ASD being able to encounter a greater number of contingencies, thereby maximising available reinforcement in their different environments and would likely increase their ability to learn from naturalistic settings (Mullins & Rincover, 1985).

### 1.3. Extinction Induced Variability

Many authors have argued that the results found in studies such as that by Pryor et al. (1969) are due to the effects of extinction. The typical behaviours the porpoises emitted were not reinforced and each time they emitted new behaviours, these were effectively placed on extinction, while the experimenters waited for "novel" behaviour in order to deliver reinforcement. It has been shown in many studies that extinction and thinning schedules of reinforcement induces variability (Eckerman & Lanson, 1969; Stokes, 1995). Taking these findings into consideration, how can we separate out the effects of extinction and schedule changes from the effects of consequential control over variability in responding?

Antonitis (1951) provided reinforcement to rats for poking their noses anywhere along a 50 cm opening, independently of location. During training the location of poking became increasingly predictable. When reinforcement was withheld, the rats showed increased variability with respect to location and now emitted responses all along the 50 cm opening. Thus, the author noted that extinction had increased variability.

Conversely, Nevin (1967) provided reinforcement to pigeons for pecking the brighter of two discs. The birds achieved 70-80% accuracy during training. When 10 extinction sessions were implemented the response rates decreased but the

accuracies were maintained at their previous levels. So, the pigeons did not vary their responses when reinforcement was withheld. This, and other research, was presented as evidence for the stable nature of responding during extinction.

Eckerman and Lanson (1969) used intermittent schedules and extinction to investigate variability of responding in pigeons. They found that variability decreased during continuous reinforcement and increased during intermittent reinforcement and extinction. However, follow-up probes were not conducted so the maintenance effects of extinction-induced variability are still in doubt. Neuringer (2002; 2004) reported that the variability elicited by a decrease in reinforcement frequency or extinction will be short-lived, while that resulting from direct reinforcement of a variability contingency, will be more resistant to extinction. This was indicated by the findings of Miller and Neuringer (2000), who showed that the increases in variability in responding demonstrated by participants with and without ASD following contingent reinforcement of variable responding were maintained in the subsequent phase when reinforcement was provided non-contingently. However, more research is needed to fully analyse maintenance effects of reinforced variability in an applied setting over time.

Balsam et al. (1998) sought to investigate whether it was the surprise element or the disappointment element of schedule change and extinction that resulted in increased variability. They taught three groups of participants to respond to a keyboard task involving a 6 x 6 square matrix. Sequences of arrow down and arrow right keys caused a light to move from the top left to the bottom right corners. At the end of every correct sequence participants received 10 points. In the next phase, Group 1 completed a block of 10 trials where they received no points, Group 2 completed 10 trials where they only received 1 point and Group 3 completed 10 trials

and received 100 points. If it was just the surprise element that produced variability then all three groups would have increased variability. However, the results showed that the two groups who had a decrease in reward showed an increase in variable responding. The group which had an increase in reward remained unchanged. This suggests that it was the reduction or removal of reinforcement that brought about increases in variability.

In a comprehensive investigation of the effect of extinction on response variability, Neuringer, Kornell and Olufs (2001) provided reinforcement to three different groups of rats who responded in sequences of three presses across three operanda (L and R levers and a key on the back wall of a chamber). In the VAR group, reinforcement was contingent on the rats emitting variable sequences. In the REP group, the rats accessed reinforcement when they emitted a single repeated sequence, while in the YOKE group the rats received reinforcement independently of variations or repetitions. Following the reinforcement phase, they implemented an extinction phase and then compared extinction after reinforcement across each of the three contingencies. Their findings were as follows:

1. Rates of responding decreased across all groups during extinction, i.e., the rats responded at a lower rate during extinction no matter what group they had been assigned to.
2. Sequences that had been emitted most frequently during the reinforcement phases were also emitted most frequently during the extinction phases.
3. Sequence variability increased across all three groups during extinction with those sequences that had been emitted most rarely during reinforcement showing the largest increase in frequency.

The authors noted the seemingly inconsistent findings that extinction resulted in the pattern of learned responses remaining unchanged while response variability increased. However, they concluded that the most common responses remained intact and the increases in variability during extinction were due to the increase in low-probability behaviours.

Furthermore, this study by Neuringer and colleagues (2001) added an important insight to this on-going debate. When they ordered the 27 possible sequences from most probable to least, there was a high correlation between those hierarchies across the reinforcement and extinction phases for all three experimental groups. They reported that this was a significant finding as it indicates that, consistent with the literature that shows maintenance of previously learned responses, previous learning was not "wiped out" (p. 91). From their data, it was clear that while high probability sequences remained intact, low probability sequences increased in frequency in a small way relative to their baseline levels and this increased the level of variability overall. This pattern of behaviour appears to be functional from an evolutionary perspective whereby organisms "stick" to behaviours that have worked for them in the past (which is functional when reinforcement is intermittent) but they implement a rule of occasionally trying something new (which is functional when reinforcement is depleted). In this way, they will be successful in times of uncertainty, and hence the common finding that extinction induces behavioural variability.

Duker and Van Lent (1991) reported that children with intellectual disabilities who had been taught to use gestures to mand often used only a small proportion of these topographies. They withheld reinforcement for high-frequency gestures and this led to an increase in different gestures being used as mands.

Similarly, Lalli, Zanolli and Wohn (1994) used extinction procedures to increase variability in toy play. At baseline, participants showed no appropriate play with toys. Firstly, the experimenters used positive reinforcement techniques to train one topography for each toy. Next, these topographies were placed on extinction and this resulted in the emergence of untrained topographies. These studies demonstrate that extinction induced variability occurs across multiple behavioural repertoires.

Rogers and Ghezzi (2010) provided reinforcement to children with ASD and neurotypical children for placing a ball into one of four holes. The target hole was highlighted with brightly-coloured paper. When the participants reached a criterion of 50 consecutive correct responses the extinction condition was implemented and there were no programmed consequences for putting the ball into any of the four holes. During extinction, the children with ASD tended to respond to a given location for an extended period of time. When they switched to a different location they would then respond there for a longer period of time before switching to the next location. The neurotypical children, however, responded for only a brief period of time (often just a couple of responses) before switching to a different location. This indicates that the neurotypical children responded more readily to the extinction component than the children with ASD as they increased they demonstrated increased variability in responding more quickly. The authors report that, in particular for children with ASD, direct reinforcement of response variability across topographies is warranted in order to produce clinically relevant increases in variability.

### 1.4. Lag Schedules of Reinforcement

Under lag reinforcement schedules a response or sequence of responses is reinforced if it differs from a specified number of previous responses (Lee, Sturmey & Fields, 2007). For example, in a lag 1 schedule a response or sequence of responses will only be reinforced if it differs from the previous one. During a lag 1 schedule, a participant could alternate between two responses throughout the reinforcement period and access reinforcement for every response. In a lag 3 schedule a response will only be reinforced if it differs from the previous three responses. Thus, greater lag schedules require greater behavioural variability to be demonstrated in order to access reinforcement.

Lee et al. (2007) described effects of different schedules of reinforcement on rates and patterns of responding. Typically, ratio schedules will produce a higher rate of responding but it will be more stereotyped while interval schedules do not produce such a high rate of responding. The authors noted that differential reinforcement of stereotyped responding is likely to occur with ratio schedules but this is not as likely with interval schedules. They also noted that variable schedules are likely to result in higher variability in responding than fixed schedules. Lag schedules may be superior to the more traditional schedules as the resultant variability is due to the schedule directly acting on operant variability, rather than variability being a by-product, as is the case with traditional schedules. The contingent reinforcement of variability, using lag schedules, allows for a strengthening of the relations between members of a class of behaviours, as opposed to the properties of a single response.

Lag schedules have been used effectively to increase response variability in human and non-human animal research (Susa & Schlinger, 2012). Page and

Neuringer (1985) demonstrated that sequence variability in the key-pecking of pigeons could be increased and maintained on a lag 50 schedule. Neuringer (2002) noted that this approximated a random response generator. Schoenfeld, Harris and Farmer (1966) used a lag 1 schedule of reinforcement to increase the variability of inter-response times in the bar-pressing of rats. The experimenters defined a series of temporal class intervals. The rats were required to vary the intervals in which their responses fell in order to access reinforcement. If a response fell in the same temporal class as the previous one then reinforcement was withheld. In this way, the experimenters were able to increase the variability in the rat's inter-response times.

Lag reinforcement schedules have also been used to increase variability in applied settings. Lee, McComas and Jawor (2002) increased variability in the verbal responses of two of their three participants with ASD using a lag 1 schedule. Susa and Schlinger (2012) noted that lag schedules of reinforcement can be arranged to teach language that more closely approximates typical conversation. They expanded the work of Lee et al. (2002) by utilising lag 1, lag 2 and lag 3 schedules to increase variability in vocal responses of a boy with ASD to a social question. They found that variability increased as they increased the lag value. Napolitano, Smith, Zarcone, Doodkin & McAdam (2010) increased variability in the colour and shape of block structures of children with ASD using a lag 1 schedule of reinforcement, while Cammileri and Hanley (2005) used a lag 1 schedule to increase variability in the selection of activities of neurotypical students who were reported have low variability in this area. More recently, applied researchers have begun to investigate the effects of increasing the lag requirement up to a lag 3, demonstrating that higher variability can be achieved when the lag value is increased (Heldt & Schlinger 2012; Susa & Schlinger, 2012). The outcomes are promising and they point to the

superiority of lag schedules versus traditional schedules when the target behaviour is variability in responding.

Lee et al. (2007) noted that a possible limitation of lag schedules is that participants may emit the minimal amount of variability required in order to fulfil the schedule requirement. This may actually lead to stereotyped patterns of responding. In particular, on a lag 1 schedule, alternating between only two responses would result in 100% of responses accessing reinforcement. When using higher lag values it is also possible that repetitive chains of responding could be shaped. For example, when using a lag 3 schedule a series of responses R1, R2, R3 emitted repeatedly would maximise all available reinforcement. This has been referred to as higher-order stereotypy (Schwatrz, 1982b) and it is a consideration for applied researchers who use lag schedules to increase variability across a variety of repertoires.

## 1.5. Measures of Variability

**1.5.1. U-Value Statistic.** Predictability of an individual's behaviour depends on both the range of alternative responses available to them and the relative frequencies at which these alternatives are emitted (Miller & Frick, 1949). A quantitative index of predictability based on these parameters is referred to by the above authors as an "index of behavioural stereotypy". The authors explain that this term was chosen to represent orderliness in behavioural sequences "without specifying whether the orderliness is due to learning, fixation or lack of understanding" (p. 317).

If there are only a few possible responses then it is easier to predict responding than if there are many alternatives. Thus, uncertainty (U) is increased when there are more alternative responses available, where the uncertainty is a

logarithmic function of the number of alternatives (Miller and Frick (1949). This can then be extended to include alternative responses of differing probabilities. The U value is a widely used statistic in the literature pertaining to operant variability (Abreu-Rodrigues, Lattal, Dos Santos & Matos, 2005; Cherot, Jones & Neuringer, 1996; Neuringer, 2002, 2004; Stokes, Mechner & Balsam, 1999,). It is calculated using the following formula (if there were 16 possible responses, i.e., sequences of four presses across two keys),

$$U = -\sum[(N_{1\text{-}16})*(\log_2[N_{1\text{-}16}])]/\log_2(16)$$

where N equals the frequency of occurrence of a given response divided by the total number of responses. The U-value looks at the distribution of response probabilities with equal probabilities indicating high variability and unequal probabilities indicating low variability (Neuringer, Kornell & Olufs, 2001). In effect, U approaches 1 when frequencies are approximately equal, and approaches 0 when one possibility predominates (Neuringer, 2004). In other words, higher U-values (closer to 1) indicate higher variability while low U-values (closer to 0) indicate lower variability or more repetitive responding.

U-values have been used in research papers to demonstrate that variability is controlled by discriminative stimuli (Denney & Neuringer, 1998); to show the effects of different acquisition procedures on response variability (Stokes, Mechner & Balsam, 1999); to investigate and modify behavioural variability in depressed students (Hopkinson & Neuringer, 2003); to compare the levels of variability resulting from decreased reinforcement versus reinforcement contingent on variable responding (Grunow & Neuringer, 2002); to examine stability and variability in extinction (Neuringer, Kornell & Olufs, 2001); to investigate the effect of lag reinforcement schedules on the variability of pigeons' sequences of responding (Page

& Neuringer, 1985) and to measure increases in variability in individuals with ASD when reinforcement was contingent upon it (Miller & Neuringer, 2000).

**1.5.2. Percentage of trials in which variability contingencies are met.** Page and Neuringer (1985) used lag schedules of reinforcement to increase variability in the sequences of key pecks of pigeons. They increased the lag requirement across the experimental phases. The measure of variability they used was the percentage of trials in which variability contingencies were met. This was calculated by dividing the number of trials that resulted in reinforcement by the total number of trials.

**1.5.3. Other Measures.** Other measures for calculating variability in responding include the frequency of novel responses emitted, conditional probability of responses, frequency of different responses in a session and statistical tests of the distribution of responses (Neuringer, 2002; Neuringer & Jenson, 2012).

## 1.6. Implications for Learning and Performance

**1.6.1. Creativity.** The language of creativity is usually reserved for the product of the behaviour (Epstein, 1996). The literature doesn't usually describe the behaviour itself. This may be due to a lack of understanding of the factors that influence creativity and variability of behaviour. Epstein (1996) refers to behaviour as being "ever-changing and ever-novel" (p. 8) and suggested that the overlapping probabilities of many different behaviours could be considered as an interconnection of repertoires. It is clear that variability in behaviour is required for such interconnections to occur. Creativity has been described as an essential component of problem-solving (Ruscio & Amabile, 1999), indicating that it is an important repertoire for successful human development.

Behavioural and cognitive psychologists have participated in much debate over the effect of reinforcement on creativity, with the latter reporting an inhibitory effect. Eisenberger and Selbst (1994) investigated these arguments and through carefully controlled experiments, they also sought to resolve some of the issues that had been raised. They noted that in studies involving reinforcement contingencies, cognitive psychologists frequently failed to use repeated presentations of reinforcing stimuli as the behavioural studies did. They also frequently reinforced low levels of creativity. These factors were determined to have affected the outcomes of their studies.

Indeed Eisenberger and Selbst (1994) report that, in many of the studies published in the behavioural literature, reinforcement of low variability was used as a reversal technique. In this way, researchers have shown that contingent reinforcement of high levels of variability results in high variability in responding. When reinforcement is contingent on low levels of variability, variable responding decreases and, when the high variability condition is reintroduced, variability increases again (Goetz & Baer, 1973).

Furthermore, Grunow and Neuringer (2002) studied four groups of rats and reinforced their sequences of lever presses across three levers. One group received reinforcement for low levels of variability, two groups for intermediate levels and one group for high levels of variability. They found that levels of variability were directly controlled by the contingencies in place.

According to Eisenberger and Selbst (1994), the behavioural studies often explained the contingencies to participants beforehand and it was difficult to separate the effects of reinforcement from the effects of instructions. While there is some support for this process, the Ross and Neuringer (2002) study described above

challenges this as the students were not given any information about the contingencies in place or what they would have to do to access reinforcement. As a result of the contingency alone, they were able to repeat along one dimension and vary along the other two. This powerful control in the absence of verbal instruction challenges the argument that the success of behavioural studies on reinforced variability or creativity is due to instructions and not the reinforcement itself.

**1.6.2. Problem-Solving.** Becker (1981) described problem-solving as "an organized effort for delivering unique and relevant solutions for opportunities and undesirable situations" (p. 2). He described it as a creative endeavour as it involves a search for unique, infrequent and relevant solutions. Problem-solving can involve a novel chain of previously-learned responses or a novel combination of previously-learned responses (Mayfield & Chase, 2002). When no instruction has been received regarding a novel problem, individuals must rely on instruction received for previous problems in order to generate a solution in the novel context. In terms of education, there has been an increasing focus on the importance of teaching and promoting general thinking, reasoning and problem-solving abilities (Bransford, Arbitman-Smith, Stein & Vye, 1985). Students need to be able to use their knowledge in flexible ways and many researchers are investigating the optimum teaching methods to enable them to do so (Ruscio & Amabile, 1999). As teachers cannot prepare their students for every problem they will encounter, it is important to make every effort to equip students with the skills to approach novel problems in a way that will facilitate creative and effective solutions. Students must be able to relate novel aspects of a problem to skills and ideas in their repertoire. In this way they can generate effective solutions.

Ruscio and Amabile (1999) compared two approaches to problem-solving that are commonly described in the psychological literature. The algorithmic approach involves following clearly-defined steps towards a specific goal. The heuristic approach involves following steps that are more general and exploratory.

Participants in the experimental groups were shown one of two videos instructing them on a structure building task. The algorithmic video showed the instructor building the structure from start to finish in a series of specific steps, similar to rote learning. Each new technique was introduced as a logical step toward the final goal. In the heuristic video, the same techniques were shown but they were loosely grouped by functional similarity and not presented in a step-wise progression. While participants in the heuristic condition were shown techniques and information that would help them to complete their task, they were not at any stage during the demonstration video shown a complete or even a partially-complete model of the structure or given information regarding how to construct it. To allow for comparison across all the groups of participants, those in the heuristic video and control (no video) groups saw a picture of the completed structure demonstrated to the algorithmic group and told it was a sample of what could be constructed. The participants then had 15 minutes to build their structure.

Results showed that both groups retained the same amount of task-related information and there was no difference between them in meeting the requirements of the task (height criterion and using all available materials). While the authors reported the relationship between instructional style and creativity, they noted that when students are taught algorithmically they will be more successful when a task is similar to one they received instruction on, but when novelty or deviation is required they perform less successfully than those taught in a more heuristic style and even

those who have not been instructed at all. Furthermore, while students in the algorithmic condition worked at a faster pace and appeared to be more confident, they were less likely to generate novel solutions to the problem-solving task than those in the heuristic condition. The authors reported that "students in the heuristic condition tended to use learned techniques in a less rigid fashion and were more likely to attempt problem-solving strategies that had not been explicitly taught to them" (Ruscio & Amabile, 1999, p. 263).

The findings of Ruscio and Amabile support those of Mayer and Greeno (1972), who also showed that algorithmic type instruction resulted in greater skill in solving problems similar to those previously mastered. However, a more heuristic type instruction led to knowledge that students could use in more flexible ways and apply to new contexts.

Mayfield and Chase (2002) compared three different methods of teaching algebra rules to college students. The training methods were the same but the groups differed in terms of the practice sessions offered during the training. The cumulative practice group received additional practice on component skills combined with practice on all skills previously mastered. The extra practice group received practice after each skill on that skill alone, while the simple review group received practice on one component skill at a time. Application tasks and tasks that involved novel combinations of the skills learned were used to evaluate differences between the three groups of students. The results showed that the cumulative practice group were significantly better than the other participants on accuracy of application and problem-solving skills as well as rate of correct performance on problem-solving skills. Their retention was also significantly better. The authors suggest that this was due to the fact that this group had received training in combining all skills

learned prior to testing. They had experience of emitting variability in their problem-solving behaviour by combining mastered skills in novel ways. The other groups did not receive this type of practice and the authors noted that they tried fewer solutions to the problems posed during the testing conditions. This study suggests that the method of combining and recombining responses was most effective for solving problems. The ability to vary behaviour and try new things is a prerequisite for this method.

An individual's ability to vary their behaviour and to try new response patterns contributes to their ability to solve problems and resolve difficult or aversive situations in their environment. Epstein (1985b) considered different performances in problem-solving tasks as a function of parameters such as the genes and history of the person, current stimuli and how they change over time as a result of the person's behaviour. "Edison, as a baby, could not invent. He had to learn to invent. Something in his environment reinforced his persistent exploratory responses" (Becker, 1981, p. 8).

**1.6.3. Verbal Behaviour.** Typical verbal behaviour is complex with a high level of variability. In contrast, children with ASD have been observed to have limited vocal verbal repertoires and they commonly use what language they have in a repetitive manner (Susa & Schlinger, 2012). Carr and Kologinsky (1983) described aspects of language such as spontaneity and flexible conversational skills as difficult to teach to individuals with ASD. This may be due to the absence of effective contingencies rather than an inherent inability of this population to vary their verbal behaviour (Miller & Neuringer, 2000). Lee et al. (2002) described how children with ASD tended to give the same response to a given question every time it was asked of them. For example, if a child answers "Fine" every time they are asked "How are

you?", this can be described as invariable intraverbal behaviour. Similarly, children with ASD have been shown to have low variability in their mand frames (Betz, Higbee, Kelley, Sellers & Pollard, 2011). They may use "I want" every time they mand. This does not reflect typical development where children will ask for things in a variety of ways e.g., "Give me" or "Can I have". This repetitive aspect of verbal behaviour in children with ASD reflects the nature of the disorder as it is defined by repetition.

The prevalence of echolalia among children with ASD is further indication of the low variability in verbal behaviour often demonstrated by this population. Rydell and Prizant (1995) report that 85% of children with ASD who acquire language, will emit echolalia. Echolalia is, by definition, restrictive and repetitive. It appears to be non-functional, though it can serve different functions for the individual such as escape from instruction. Children who emit echolalia may not come into contact with contingencies that select and maintain variability in language. If they are not displaying the variable foundation that Skinner (1938) refers to then they will not have the opportunity to receive reinforcement for such variability. This could mean that this group of children would retain low variability in verbal behaviour into adulthood.

In one of the earliest investigations into methods for increasing variability in the verbal behaviour of children with ASD, Lee, McComas and Jawor (2002) provided reinforcement to three boys with ASD for responding variably to a social question using a lag 1 schedule but the responses had to be appropriate to the context. They selected questions for which the participants typically gave the same response. For two of the participants the question was "What do you like to do?", while for the third participant the question was "How are you?" They noted increases

in appropriate variable responses for two of their three participants. The same two participants also demonstrated increases in the cumulative number of novel responses to the social question across the study period. The authors noted that the participants who answered the question “What do you like to do?” often referenced items that they observed in their environment when replying to the question. They reported that it was likely that the response was functioning as a mand for those items in some of the trials, and that the uncontrolled variation of stimuli in the environment may have prompted the variation in responding of these participants.

Lee and Sturmey (2006) attempted to control for the variables in the previous study by asking the same social question of their three participants with ASD (“What do you like to do?”) and by systematically controlling the number of preferred items present in the course of the intervention in a reversal (ABA) design. They conducted a preference assessment to identify 10 preferred and 10 non-preferred items for each participant. At the beginning of each session a total of 10 items were placed on a desk next to the experimenter. Across three different conditions either 10, five or none of these items were preferred, i.e., 100%, 50% or 0% of the items present were preferred. Following a baseline phase where appropriate responses were reinforced on a lag 0 schedule, a lag 1 schedule was introduced. Results showed that, similar to the previous study, response variability increased for two of the three participants in the presence of the lag 1 schedule. Furthermore, the proportion of preferred stimuli present in the environment did not affect the level of variability demonstrated for any of the three participants in this study.

Susa & Schlinger (2012) further expanded the work of Lee and colleagues by utilising lag schedules of reinforcement to increase the number of different responses required of a boy with ASD before he accessed reinforcement while systematically

increasing the lag requirement in a changing criterion design (lag 1, lag 2 and lag 3). The target behaviour was variability in verbal responses to the social question "How are you?" as measured by the average number of previous responses from which a given response varied. They observed that the number of responses from which each response varied increased steadily across the lag 1, lag 2 and lag 3 conditions. Heldt and Schlinger (2012) sought to increase variability in tact responses, using a lag 3 schedule of reinforcement with two boys, one of whom had mild intellectual disability, the other of whom had Fragile X syndrome. They sought to expand the previous work by including follow-up probes to assess for maintenance effects of the procedure. Unlike Susa and Schlinger's study, here the lag 3 schedule was implemented directly following the baseline phase. They found that variability in tacting increased for both participants and these gains were maintained at a three-week follow up. The authors posit that the reason for the maintenance effect in their study is likely to be that the lag 3 procedure more closely approximates the contingencies encountered in the natural environment than the lag 1 procedures employed by Lee and colleagues, who observed a return to baseline variability once the schedule was removed.

Jones (1990) noted that when individuals exhibit repetitive or stereotyped responding to social questions it can result in stigmatisation, and restriction of opportunities for more complex verbal interactions. As more children with ASD and other developmental disabilities are being integrated into mainstream educational and recreational settings, a greater degree of variability across the verbal operants is of increasing importance to this cohort.

**1.6.4. Play.** Play has been shown to lay the foundation for adaptive functioning in later life (Pellegrini, Dupuis & Smith, 2007; Wenner, 2009).

According to Sutton-Smith (1975), play behaviour in early years is repetitive but becomes more complex. It is where a child develops the capacity to deal with novel situations and behave in variable ways when their culture demands it. Through play, children learn to respond to a variety of discriminative stimuli. Play encourages flexibility and creativity that is likely to benefit the individual throughout their lifespan.

Spinka, Newberry and Beckoff (2001) proposed a hypothesis of mammalian play that they termed “training for the unexpected”. In this hypothesis, play is viewed as a way to rehearse the behavioural sequences where some motor control or balance is impaired and the animal must react quickly in order to survive. Similarly, they must learn to react quickly and effectively in novel situations; their play prepares them for this from a young age. The authors also noted that higher levels of play occur among species that regularly experience changing habitats. As these species experience greater unpredictability in their environment, they stand to benefit from rehearsing the skills needed to establish routes through their territory. Through play, animals learn to respond to a variety of unfamiliar situations. Animals that can use a variety of skills, even in novel situations have the greatest chance of survival. Driver and Humphries (1988) investigated unpredictability in the behaviour of animals and found that survival was the primary function. Across many different species, they noted variability in speed, topography of movement and unpredictable changes in direction.

According to Wenner (2009) it is the variable nature of free play that most contributes to adaptive functioning in later life. It fosters creativity and has implications for the individual’s functioning in the domains of social, emotional and cognitive development. Pellegrini (1995) examined rough and tumble play in young

boys and found that the ability of the boys to switch from one activity to another during this type of play was highly correlated with their scores on a test of social problem-solving ability in later years. This suggests that the skills developed through play result in greater flexibility in behaviour in response to novel circumstances in later life. "Juveniles use play to sample their environment in order to develop adaptive behaviours….play influences the process of [evolution] by supporting the development of new strategies in novel environments during the juvenile period" (Pellegrini, Dupuis & Smith, 2007, p. 261).

Neuringer (2002) suggested that the self-maintaining nature of play serves as evidence that engaging in variable behaviour may be reinforcing for the individual. The abundance of research demonstrating that varying reinforcers increases the value of those reinforcers (Bowman, Piazza, Fisher, Hagopian & Kogan, 1997) may provide further support for this hypothesis. In the case of individuals who demonstrate low variability, reinforcing variability and conditioning a variety of toys, may have important positive implications for future development.

Goetz and Baer (1973) delivered verbal praise to pre-school children each time they made new patterns using building blocks. The authors noted that novel patterns were continually produced by the children. During a reversal condition, the praise was contingent on repeating previously made structures and novelty decreased. However, when novel patterns were again reinforced the variety of new structures again increased.

Napolitano, Smith, Zarcone, Doodkin and McAdam (2010) extended this work in an attempt to increase diversity in the block-building of six participants with ASD. However, instead of reinforcing novel responses they delivered reinforcement if a response differed from the prior response, i.e., they delivered reinforcement on a

lag 1 schedule. The lag schedule was employed for two of the participants for varying the colour of blocks used, and for the other four participants for varying the form of the structures. All participants increased their response diversity, although four of the six participants required teaching sessions where the experimenter modelled variant responses in order to achieve this aim. Two of the participants failed to return to baseline rates in the second baseline condition. The authors noted that this may be as a result of the new diversity in responding acquiring reinforcing value.

The findings of Goetz and Baer (1973) and Napolitano et al. (2010) are important as they demonstrate that variability in play behaviour can be increased using reinforcement procedures. However, the target behaviour under study, namely block-building is an early play skill. In contrast, it may be more difficult to increase variability across the more complex levels of play, e.g., functional play and symbolic play.

Different and increasingly complex levels of play are observed in the course of typical development (Libby, Powell, Messer & Jordan, 1998). Sensorimotor play emerges in the first year of life and involves behaviours such as banging, mouthing and spinning. Here, the infant does not consider the functional properties of the object or toy. Following this, the child begins to emit relational play where objects are combined, e.g., by being placed inside each other or stacked on top of each other.

The next stage, functional play is described as the use of an object in a conventional way or the conventional combination of two objects such as using a spoon to feed a doll (Ungerer & Sigman, 1981). Finally, symbolic play has commonly been deconstructed into three different forms (Jarold, Boucher & Smith, 1993; Leslie, 1987):

1. Object Substitution, which is described as using one object to represent another, for example, using a block as a bar of soap;
2. Attribution of Absent or False Properties to an item, for example, pretending a dry table is wet;
3. Imaginary Objects Present, for example, pretending an empty vessel contains water.

It is clear that symbolic play involves a high level of variability. Methods for increasing variability across the different levels of play in the clinical population may assist greatly with social interaction and their integration into inclusive settings.

### 1.7. Generalisation of Variability to Untargeted Repertoires

Goetz (1982) reported that the generalisation effects of reinforced variability are more pronounced when they fall within similar response classes. Thus, when measuring effects of variability training, a greater effect will be seen when the transfer task is similar to the training task. This suggests that choosing tasks from a variety of repertoires for variability training will improve outcomes and result in greater overall variability in behaviour than targeting specific or narrow repertoires.

Maltzman (1960) repeatedly presented college students with a list of stimulus words and asked them to produce free associations for each stimulus word. Reinforcement was provided for emitting a new word each time a stimulus word was presented. When he then asked the students to give uses for an array of common objects, these students generated significantly more unusual uses than a control group. While this study demonstrates a generalisation effect for variability in responding, the tasks were quite similar and further research is necessary to

investigate generalisation across repertoires following reinforcement of variable responding.

Holman, Goetz and Baer (1977) provided reinforcement to preschool children for novel drawings and found that novelty increased as a result. When they observed the same children building block constructions they demonstrated some transfer of novel responding. In the subsequent phase, children of the same age engaged in four tasks independently: painting, drawing with a pen, Lego® and wooden blocks. When the children received reinforcement contingent upon variability in their pen drawings, they also demonstrated variability in their paintings – a topographically similar task. However, they did not transfer this change in variability to the topographically different Lego or block building tasks. While there is evidence that operant variability can transfer to new skills, the extent to which generalisation can occur across repertoires is unclear. Napolitano et al. (2010) increased variability in the block building of six children with ASD using a lag 1 schedule of reinforcement with colourful plastic blocks. They conducted generalisation probes with wooden blocks of similar size and colour but only one participant demonstrated a generalisation effect of variability to the new stimuli, despite the similarity to the training stimuli. Five of the participants had at least moderate problems with restricted or sameness behaviour as indicated by the Repetitive Behaviour Scale – Revised (Bodfish, Symons, Parker & Lewis, 2000). Participant characteristics may be an important variable in predicting how successful students will be at generalising increases in variability to new stimuli and situations.

Following the lag reinforcement schedule which increased variability in vocal responding to a social question, Lee et al. (2002) investigated whether this variability would generalise to new settings and listeners. During generalisation sessions any

appropriate response to the social question was reinforced, independent of variability. They found that varied responding generalised to different therapists and settings in the absence of the lag schedule.

Harding, Wacker, Berg, Rick and Lee (2004) used contingent reinforcement of novel responses to increase variability in the kicking and punching techniques of two martial arts students. They implemented the intervention in drill-type training sessions. Next, they conducted a more realistic sparring condition that resembled a naturalistic context for the demonstration of technical skills. They found that the increase in variability generalised to the untrained context.

Neuringer (2002) noted that the extent to which transfer occurs is one of the most important of the remaining unresolved questions in the area of operant variability. It is clear that there is a need for research in this area to further understand how generalisation of response variability to new repertoires may occur and to identify the optimum methods for promoting this transfer.

### 1.8. Evidence of Response Variability in ASD

Repetitive and stereotyped behaviours are characteristic of individuals with ASD. This may be regarded as a lack of behavioural variability as such behaviours are, by definition, invariable. It has been demonstrated through research that children with ASD are prone to perseveration (Mullins & Rincover, 1985); do not sample available options in their environment (Boucher, 1977); often respond to a limited portion of their environment (Lovaas, Schreibman, Koegel & Rehm, 1971), and frequently demonstrate repetitive body movements and object manipulations, insistence on sameness of environment or daily routines and circumscribed interests (Gabriels et al., 2008).

For example, Boucher (1977) asked children with ASD and neurotypical controls to choose one of two arms of a maze. The neurotypical children varied between the two arms, while the children with ASD generally chose only one arm. When a third arm was added, the neurotypical children were significantly more likely to follow it than the children with ASD i.e., the children with ASD ignored the new option, suggesting a deficiency in sampling behaviour. Miller and Neuringer (2000) provided reinforcement contingent on, and independent of, variability in button pressing across different experimental conditions and found that individuals with ASD responded less variably overall than adult and child controls.

Baron-Cohen (1992) asked neurotypical children, children with intellectual disability who did not have ASD and children with ASD to hide a penny in one of their hands so that the experimenter would not be able to guess which hand it was in. He found that children with ASD were more likely than their peers in the other two groups to generate a predictable pattern such as switching between right and left hands repeatedly and they frequently perseverated on one hand. The children with intellectual disability didn't perseverate on one hand but they typically switched between two hands, demonstrating some impairment in variability. The neurotypical children were able to produce more random patterns so that it was more difficult for the experimenter to guess which hand held the penny.

Mullins and Rincover (1985) noted that one of the most basic skills in animals and humans is the ability to maximise reinforcement. When faced with concurrent variable interval schedules of reinforcement, organisms will respond in a predictable way and follow the matching law which states that "the relative frequency of responding on an alternative matches the relative frequency of reinforcement on that alternative" (Mullins & Rincover, 1985, p. 352). When faced

with two variable ratio schedules, maximising occurs when the response with the most lucrative reward/cost ratio is selected exclusively. Children with ASD may not be able to match or maximise reinforcement when faced with concurrent schedules (Rincover & Koegel, 1975). This may be due to low operant variability in this population. They may be stuck on one alternative because they simply have not attempted to sample another one.

In order to assess the ability of children with ASD to maximise reinforcement and sample available stimuli, Mullins and Rincover (1985) observed the responses of six children with the disorder to five concurrent schedules of reinforcement. They compared them to one group of neurotypical controls matched for chronological age and one group of neurotypical controls matched for mental age.

During the training phase, five cards with a different Greek letter embossed on each were presented simultaneously and participants were asked to pick a card. Prompts were used in this phase and participants received reinforcement for removing a stimulus card from the display until each of the stimuli had been reinforced once. Next, the differential reinforcement phase was implemented. Here, each of five stimulus cards was assigned a different schedule of reinforcement (CRF, FR2, FR4, FR7, FR11). Participants were instructed to pick a card and this phase continued until the participant selected a particular card 80% of the time across two consecutive blocks of ten trials. Once this criterion was reached, the selected stimulus was put on extinction while the other four cards remained unchanged. If a child was perseverating on this stimulus card after 100 trials, it was replaced with a blank white card to encourage sampling of other stimuli. Training continued until the 80% criterion was achieved on a second stimulus card.

Each of the neurotypical participants reached criterion on the richest schedule (CRF), therefore maximising reinforcement. Only one out of the six participants with ASD did so. Of the other five, three met criterion on the FR2 schedule and two on the FR7 schedule. The neurotypical participants all maximised reinforcement again by reaching criterion on the FR2 schedule in the second phase. Five out of the six children with ASD again passed up the CRF schedule, choosing from the stimuli matched to thinner schedules. On determining how many trials it took the children to reach criterion (i.e., how efficient they were), the children with ASD took significantly longer to reach criterion than the neurotypical children.

The ASD group sampled fewer of the stimuli and perseverated on the stimulus card that was on extinction for longer than any of the control groups. This research provides strong evidence for the hypothesis that individuals with ASD are not as effective at sampling stimuli and choosing those which will maximise reinforcement for them. This is likely to be related to a lack of variability. In the first phase of this research one participant sampled only one stimulus card in the whole phase and reached criterion on that card. When this is generalised to natural settings it implies that this population will have difficulty sampling novel stimuli in the world around them and therefore may miss out on many opportunities to access a variety of social and tangible reinforcers.

The authors note that the participants with ASD demonstrated high rates of "repetitive, stereotyped motor movements such as rhythmic rocking or hand-flapping." (p. 353). This study concludes that the participants had an inability to emit variable responses and so are unlikely to sample the available alternatives and maximise reinforcement on concurrent reinforcement schedules.

The goal of habilitation as it relates to applied behaviour analysis (ABA; Cooper, Heron & Heward, 2007) is particularly relevant here. If practitioners are to have habilitation as a goal of behavioural intervention then it seems clear that reinforcement of variability in behaviour must be a component of all ABA programs for individuals with ASD in particular. Increasing operant variability in repertoires such as exploration, verbal behaviour and imaginative play may have significant and wide-reaching implications for the success of behavioural interventions.

### 1.9. Restricted Repetitive Behaviours and Low Variability in ASD

The DSM-IV (American Psychiatric Association, 2000) cites repetitive behaviour as one of the three defining features of ASD. However, compared with the extensive literature published on the other impairments related with the disorder, this class of behaviours continues to be relatively underrepresented in the literature (Lam, 2004). This fact is especially salient when one considers the challenges that repetitive or stereotyped behaviours can present for individuals with ASD, their families and service providers (Turner, 1999) and the diagnostic significance of such behaviours (Bodfish, Symons, Parker & Lewis, 2000).

Much of the early literature on restricted repetitive behaviours (RRBs) in ASD and other developmental disabilities tended to focus on stereotypical behaviours such as rocking or hand-flapping. These behaviours were considered to be defined by movement invariance and movement repetition (Rapp & Vollmer, 2005). Although repetitive behaviour was initially thought to have no obvious function (Turner, 1999), it is thought now that the main function is automatic reinforcement (Rapp & Vollmer, 2005) although it has been shown to serve a variety of functions for different individuals e.g., social positive and negative reinforcement.

In the course of their work on developing a scale to measure RRBs, Bodfish, Symons, Parker and Lewis (2000) identified the following six subscales: stereotyped behaviour, self-injurious behaviour, compulsive behaviour, ritualistic behaviour, sameness behaviour and restricted behaviour. All of these can be described as low variability behaviours.

An in-depth analysis of clinical features associated with the presence of RRBs in individuals with ASD was conducted by Gabriels, Curraco, Hill, Ivers & Goldson (2005). These authors found that the low non-verbal IQ individuals had a significantly greater presence of RRBs than the high non-verbal IQ group, thus, low variability in behaviour was related to lower IQ. They also noted that lower adaptive ability was strongly related to the presence of higher rates of RRBs. Furthermore, there was a significant correlation between greater sleep problems and higher rates of RRBs and the authors also reported significant positive correlations between rates of repetitive behaviour and the three subscales of the Aberrant Behaviour Checklist (Irritability, Lethargy and Hyperactivity). This research indicates additional negative impacts of low variability in behaviour.

Napolitano et al. (2010) reported that the findings from the experimental analysis of behaviour relating to systematically reinforcing variability in responding could offer a promising alternative to more traditional behaviour reduction-based procedures used with individuals who exhibit RRBs. Neuringer (2002) also noted that providing reinforcement contingent on variability in behaviour may help to decrease such ritualistic and stereotyped behaviours. In this way, direct reinforcement of variability in behaviour could become the intervention of choice for individuals who demonstrate high rates of RRBs. This has clear practical and ethical

implications as positive, reinforcement-based interventions are always favoured over punishment or extinction-based procedures.

### 1.10. Play and Variability in ASD

Children with ASD have been shown to demonstrate repetitive and rigid patterns of play, i.e., they have low variability in their play behaviour (Atlas, 1990; Stahmer & Schreibman, 1992). The effects of this are likely to be two-fold. On the one hand they are missing the chance to accumulate the complex patterns of behaviour that allow them to develop the capacity to deal with novel situations. Furthermore, this type of behaviour can result in an unwillingness of their peers to play with them, not just because they may have impaired social and communication skills, but because the peers may quickly become bored by repetitive play.

In addition, while children develop skills during play, it is necessary for them to emit some variations in their behaviour in order to access the contingencies that strengthen these, namely, parental approval, peer interactions and the intrinsic enjoyment of the activity. It follows that if a child with ASD does not have the ability to vary their behaviour sufficiently to access the naturally occurring reinforcers in the environment, then shaping variability in play through direct reinforcement of variability may be advantageous. If programs for children with ASD include direct reinforcement of variability in play behaviours, it is likely that their adaptive functioning in later life will be superior to those who do not learn this skill.

Gorlitz (1987) reported that early exploratory behaviour was important as this was how children become familiar with objects. He noted that this type of behaviour was correlated with the emergence of pretend play in some children. Royeurs and van Berkelaer-Onnes (1994) built on this early work and proposed that it was the

lack of exploratory behaviour in ASD that resulted in a deficit in pretend play. Thus, the lack of variability in exploration of children with ASD demonstrated by Mullins and Rincover (1985) may be related to the lack of variability in toy play shown by these individuals.

The type of play behaviours exhibited by children with ASD has been the subject of much debate and discussion. It is widely documented that children with ASD engaged in rigid and repetitive behaviours with toys (Atlas, 1990). They can often show significant deficits in independent toy play, mainly due to repetitive and stereotyped behaviours and attachment to specific toys. This can impact negatively on sampling and engagement with other toys (Paterson & Arco, 2007). As independent play acts as a precursor to the social types of play, namely, parallel play and reciprocal or interactive play, remediating these difficulties at the level of independent play may have important implications for the development of more complex play and social skills.

Many researchers report that the main difficulties for children with ASD lie in the areas of functional and symbolic play (Libby et al., 1998). Furthermore, there are disagreements among researchers with some reporting deficits in functional and symbolic play (e.g., Singurer & Ugman, 1994), while others report deficits in symbolic play only (e.g., Baron-Cohen, 1987).

Jarold, Boucher and Smith (1993) noted that there was evidence that children with ASD could demonstrate symbolic play. However, they differentiated between true symbolic play and stereotyped symbolic play, the latter being more common in children with ASD. This type of play may be as a result of particular play actions being taught by adults and is invariable in nature. By utilising procedures to increase variability in play it is likely that more “true” symbolic play could be seen in this

population. Mundy, Sigman, Ungerer and Sherman (1986) compared the play of a group of children with ASD, a group of children with intellectual disability and a group of neurotypical children. They reported that the children with ASD demonstrated fewer different functional and symbolic play acts than the other two groups. In light of recent research on operant variability, this outcome may be a result of the children with ASD demonstrating decreased variability in their play behaviours relative to their peers, rather than an inherent inability to produce functional and symbolic play.

A variety of techniques grounded in the principles of Applied Behaviour Analysis have been developed to improve and increase the play behaviours of children with ASD (Stahmer, Ingersoll & Carter, 2003). There is strong support for the use of discrete trial methods (Smith, 2001), pivotal response training (Koegel, O'Dell & Koegel, 1987) and video modelling (Paterson & Arco, 2007), to name a few. However, such procedures do not directly target what has been shown to be a significant deficit for this population, variability in sampling of toys and in play actions with and without toys.

While there are a few recent studies emerging that have demonstrated methods for increasing variability in the verbal behaviour of children with ASD, the same cannot be said for variability in play. As described above, Napolitano et al. (2010) implemented a lag 1 schedule with six children with ASD to increase variability in the colour of blocks and structural forms during a block-building task. For four of the participants teaching trials were required, as the lag 1 schedule alone, did not increase variability. However, all six participants demonstrated increased variability across the intervention period, demonstrating that the package which included a lag schedule was successful at increasing variability in this task.

### 1.11. Summary

The above review outlines the origins of research into operant variability in the Experimental Analysis of Behaviour. This field of research continues to produce new information regarding the nature of variability and optimum conditions for promoting it. The idea of reinforcing variable patterns of responding in the clinical population is not a new one but it has been slow to permeate the literature. This is surprising, given the necessity of behaving variably in order to succeed in our world. This may be due to the fact that as infants, the majority of individuals readily learn to vary their behaviour by encountering numerous environmental contingencies. Given the social importance and value of creativity and problem-solving skills, and of varying behaviour across the repertoires of play, language and social skills, these frequently are selected and maintained through social reinforcement. However, problems can arise when individuals do not develop the ability to vary their behaviour e.g., many individuals with ASD diagnoses.

The techniques that arise from human and non-human animal research will allow us to teach variability through direct reinforcement. This may act as a pre-requisite skill for other higher-order behaviours. Mullins and Rincover (1985) discuss the keystone deficits in ASD. By developing strategies to remediate these, we may see vast improvements in other skills without them being directly targeted. To this end, the emergence of several applied studies targeting increases in operant variability, including verbal behaviour and play, in recent months is encouraging.

Behavioural variability is related to greater adaptive functioning due to the ability to adapt and apply skills in new ways and in different environments. It is an essential component of creativity, problem-solving, verbal and non-verbal

communication, meaningful independent and social play and exploration of environment to take advantage of available reinforcement. It is valued in the workplace and in academic pursuits as well as in social contexts. Thus, the identification of efficient and effective methods for increasing variability is an important pursuit that has potential to improve outcomes among the clinical population.

# Chapter 2

## Study 1

## Response Variability in Children with Autism Spectrum Disorder: Comparisons with Neurotypical Children and Relationships to Restricted Repetitive Behaviour Subtypes

## 2.1. Introduction

Researchers in the field of the Experimental Analysis of Behaviour (EAB) have been investigating behavioural variability for several decades (e.g., Blough, 1966; Maltzman, 1960; Pryor, Haag & O'Reilly, 1969). In this time they have provided vigorous evidence for the operant nature of variability (Neuringer, 2002). Furthermore, they have devised robust methods for increasing variability, one of which is using lag schedules of reinforcement (Page & Neuringer, 1985). As more applied researchers begin to investigate these methods, growing discussions of the merit of using such procedures to decrease restricted repetitive behaviours (RRBs) in the clinical population, has been observed (Napolitano, Smith, Zarcone, Doodkin & McAdam, 2010).

The term restricted repetitive behaviours (RRBs) refers to a broad class of behaviours characterised by repetition, rigidity, and invariance (Lewis, Tanimura, Lee & Bodfish, 2007). Lam, Bodfish and Piven (2008) reported that numerous studies had provided an explanation of RRBs involving two factors, a "lower order" category called Repetitive Sensory and Motor Behaviours (RSMB) and a "higher order" category called Insistence on Sameness (IS). While the RSMB category are observed across different disorders (e.g., Tourette's syndrome), the IS category may be especially important in autism research and intervention as it appears to be particularly characteristic of the disorder.

Lam et al. (2008) carried out an exploratory factor analysis using the scores of 316 participants with autism spectrum disorder (ASD) on 10 items of the Autism Diagnostic Interview – Revised (Lord, Rutter & Le Couteur, 1994) that assesses restricted repetitive behaviours. They recommended that a third category or factor, Circumscribed Interests (CI) be added as it appeared to be distinct from RSMB and

IS. In their sample, Lam et al. (2008) noted that RSMB was more prevalent in younger participants, those with lower IQ and those with greater social and communication deficits. IS was related to social and communication deficits and possibly greater autism severity while CI was not significantly related to IQ or any of the other autism symptom domains on the ADI-R.

Despite the fact that the presence of RRBs is one of the diagnostic features of autism (American Psychological Association, 2000), to date these characteristics have not received the same level of attention in the research literature as the social and communication deficits (Turner, 1999). Furthermore, they form a class of behaviour that many clinicians and researchers consider a problematic and pervasive aberrant behaviour to target with intervention (Cunningham & Schreibman, 2008).

With applied researchers recommending that methods derived from the research on operant variability be investigated as remediation strategies for RRBs, this represents a shift in the conceptualisation of RRBs from a problem behaviour that should be targeted through behaviour reduction procedures (Rapp & Vollmer, 2005), to a deficit that may be addressed through reinforcement strategies (Lee, Sturmey & Fields, 2007). The relationship between RRBs and behavioural variability has not been the subject of investigation to date. While researchers have shown that individuals with ASD demonstrate lower variability than their neurotypical peers (Baron-Cohen, 1992; Miller & Neuringer, 2000), the level and type of RRBs associated with low variability has not been investigated. Miller and Neuringer (2000) investigated variability in responding on a computer task with adolescents with ASD (n=5) and child (n=4) and adult (n=5) controls. They demonstrated that the adolescents with ASD showed lower variability than the other two groups.

In 1992, Baron-Cohen investigated response patterns and deception in individuals with ASD using a simple game, commonly observed in parent-child interactions. He reported that the children with ASD (n=15) were more likely to generate repetitive and predictable patterns than those with intellectual disability (n=15) or their neurotypical peers (n=15). However, neither of these studies measured RRBs or their relationship to low variability in participants with ASD. Thus, the extent of the relationship between patterns of variability in responding and the presence, severity and type of RRBS in individuals with ASD is unknown.

The "Penny-Hiding Game" (Baron-Cohen, 1992) was implemented in the current study as a measure of response variability. The first aim was to investigate the difference in levels of variability shown by participants with and without ASD using a larger sample than Baron-Cohen (1992) and Miller and Neuringer (2000). The second aim was to determine whether lower behavioural variability was related to higher rates of RRBs among the participants with ASD. The third aim was to determine whether particular categories of RRBs would be more or less related to low variability among this population.

## 2.2. Method

### Participants

Two groups of children participated in this study. Group 1 consisted of 65 children with a diagnosis of ASD (56 male, 9 female). Group 2 consisted of 65 neurotypical children (27 male, 38 female). The mean age of the participants in Group 1 was 107.6 months (SD: 37.15, range: 49 months to 203 months). The mean age of the participants in Group 2 was 112.28 months (SD: 42.63, range: 48 months to 199 months). All participants in Group 1 had obtained a diagnosis of ASD in line

with DSM-IV criteria, from a registered psychologist, who was independent of this study. Participants were recruited from mainstream schools and special schools and through personal contacts.

**Settings**

The test was predominantly conducted in participants' schools. Children were brought to quiet areas of their classrooms or else a chair was brought outside the door to the hallway. For some of the neurotypical participants it was not possible to visit their school so the experimenter visited their home instead. In this case, the test was conducted in a quiet area of their house. The test took approximately two minutes to administer. Additionally, the RBS-R was required for each participant with ASD. These were either completed on the day or returned by post within one week.

**Measures**

The Repetitive Behaviour Scale-Revised (Bodfish, Symons & Lewis, 1999) is a 43-item scale where the items are rated on a four point Likert scale from 0 (behaviour does not occur) to 3 (behaviour occurs and is a serious problem). The items are grouped into six subscales: Stereotyped Behavior, Self-Injurious Behaviour, Compulsive Behaviour, Ritualistic Behaviour, Sameness Behaviour and Restricted Behaviour. Lam and Aman (2007) conducted an exploratory factor analysis to test this structure using a large sample (307 individuals with ASD). They proposed a five-factor solution as follows: Ritualistic/Sameness Behaviour, Self-Injurious Behaviour, Stereotypic Behaviour, Compulsive Behaviour and Restricted Interests. Collectively, the five-factor structure accounted for 45% of the variance.

The factors achieved Chronbach's alpha scores ranging from 0.78 to 0.91, showing acceptable internal validity. The inter-rater reliability was 0.88 and test-retest reliability was 0.71. The Lam and Aman (2007) scoring solution was used in this study.

**Procedure**

The RBS-R was completed by the teachers, special needs assistants or tutors of the children in Group 1. Staff members completing the questionnaire were required to work directly with the participant on a daily basis and have done so for at least six months.

Next, the Penny-Hiding Game (Baron-Cohen, 1992; Oswald and Ollendick, 1989) was carried out with each participant individually. Baron-Cohen conducted the game in two conditions with 12 trials in each condition. However, he used the first six trials in each condition for scoring. In the current study, Condition 1 was identical to that described by Baron-Cohen but Condition 2 only contained six trials.

**Condition 1**. At the start of Condition 1, the experimenter showed the participant a coin and said "now I'm going to hide this coin in one of my hands and you must guess which hand it's in". There were three phases in this condition, A, B, and C. Phase A was referred to as the "no-lose" phase. Here, the experimenter actually had a coin in each hand so the participant always perceived that they had guessed correctly. The experimenter provided lots of encouragement using phrases such as "I can't believe you beat me again" and "you're just too good at this game". This "no-lose" phase consisted of four trials. One trial was a sequence of one person hiding the coin and one person choosing a hand.

Phase B was referred to as the “no-win” phase. Here, the experimenter showed the participant a coin and then discretely dropped it into her back pocket so that there was no coin in either hand. When the participant guessed “incorrectly” the experiment said “I got you that time” or “I think I’m getting better at this”. This phase also had four trials and the purpose was to motivate the child to keep engaging with the game and to try to get back to their prior winning streak.

Phase C was the “real game” phase where the experimenter played competitively by hiding one coin unpredictably. This phase also had four trials.

**Condition 2.** In this condition, the child was asked to hide the coin so that the experimenter could not find it. This condition consisted of six trials. If the participants played predictably (kept the coin in the same hand or switched hands for every trial) the experimenter guessed incorrectly for 50% of the trials so that the participant experienced “winning” and “losing”.

## Scoring

A scoring system was devised based on Baron-Cohen’s (1992) system. His original scoring included an additional measure that took into account whether the participants engaged in “cheating” or other deception techniques that would suggest the presence of a theory of mind in the participants. However, as the current study was concerned with the variability demonstrated by the participants in selecting a hand of the experimenter (Condition 1) and in choosing which of their own hands to hide the coin in (Condition 2), the additional measure was not required.

The current coding system was devised in order to determine patterns of responding demonstrated by the groups. Firstly, the data was screened and a ranked variability system was developed. Each data set was scored in three parts, each

containing six trials. For each part, a code was assigned according to the strategies used by the participants. Repeating (R) was defined as choosing the same hand repeatedly for at least four trials in a row out of the six trials (e.g., LLLL). Switching (S) was defined as any four trials where the participants switched hands alternately (e.g., RLRL) or using an Irregular Strategy (I) which was any sequence of responses that did not meet the criteria for R or S (e.g., RLRR).

Categorising the data into three parts with three potential codes for each part resulted in 10 possible variability patterns for each participant. Table 1 lists the 10 patterns, ordered from least variable (RRR) to most variable (III), along with the numerical score that was assigned to each one. A score was applied to each participant according to the level of variability they demonstrated across the 18 trials of the game. A score of 1 indicated least variable patterns of responding and a score of 10 indicated most variable patterns.

| Variability Pattern | **Score** |
|---|---|
| RRR | 1 |
| RRS | 2 |
| RSS | 3 |
| SSS | 4 |
| RRI | 5 |
| RSI | 6 |
| SSI | 7 |
| RII | 8 |
| SII | 9 |
| III | 10 |

*Table 2.1:* Variability patterns and associated codes as outcomes of the Penny-Hiding Game.

**Data Analysis**

The number of participants in each group who achieved each variability score was tallied in order to produce a graphical representation of the obtained scores. The two groups were compared to determine if there was a significant difference in variability in responding between them using a Mann-Whitney U Test. A transformation was performed on the variability scores to stabilise the variance in the residuals according to the following formula: $y = \sqrt{10-x}$, where x was the original value. Next, a *t*-test was conducted on the transformed scores to ensure that the outcomes would be robust with regard to statistical method.

Variability scores for Group 1 were compared to their scores on the RBS-R, including scores on each subtest, in order to identify if restricted repetitive

behaviours were correlated with low or high variability and to investigate whether scores on a particular RBS-R subtype were more or less highly correlated with low variability than others.

## 2.3. Results

### Descriptive Statistics

The mean variability score for Group 1 was 5.23 (SD = 2.67). The median score was 5 and the range was from 1 to 10. The mean variability score for Group 2 was 8 (SD = 2). The median score was 9, while the range was from 3 to 10. In Group 1, nine participants achieved the lowest possible variability score (1) as they demonstrated the pattern RRR. This represents a high level of repetition across the test. No participant from Group 2 demonstrated this pattern of responding. Conversely, 16 participants from Group 2 demonstrated the highest possible variability score (10), demonstrating the pattern III. Only four participants from Group 1 demonstrated this pattern. Figures 2.1 and 2.2 represent the number of participants in Group 1 and Group 2 respectively who achieved each of the 10 variability scores.

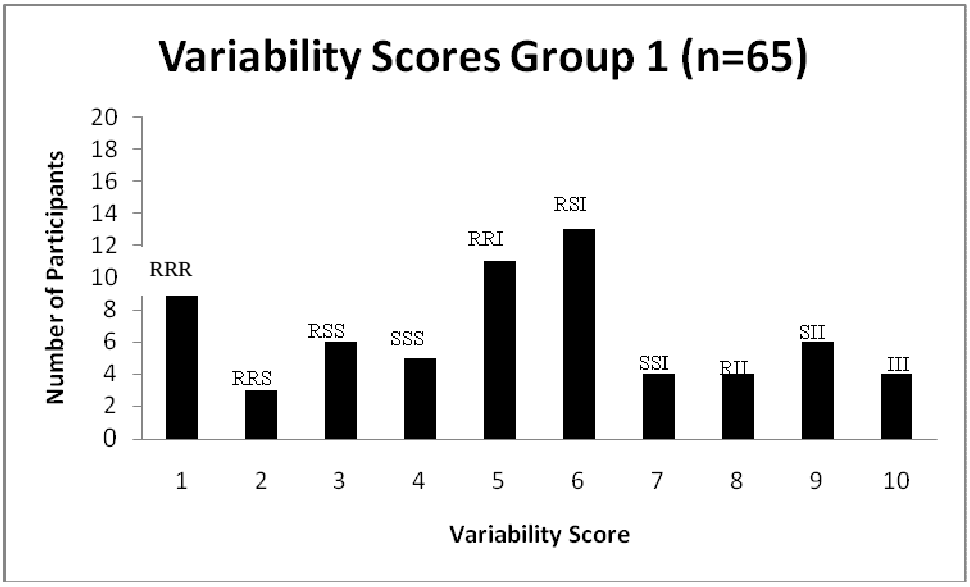


*Figure 2.1:* Number of participants in Group 1 achieving each variability score (1-10).

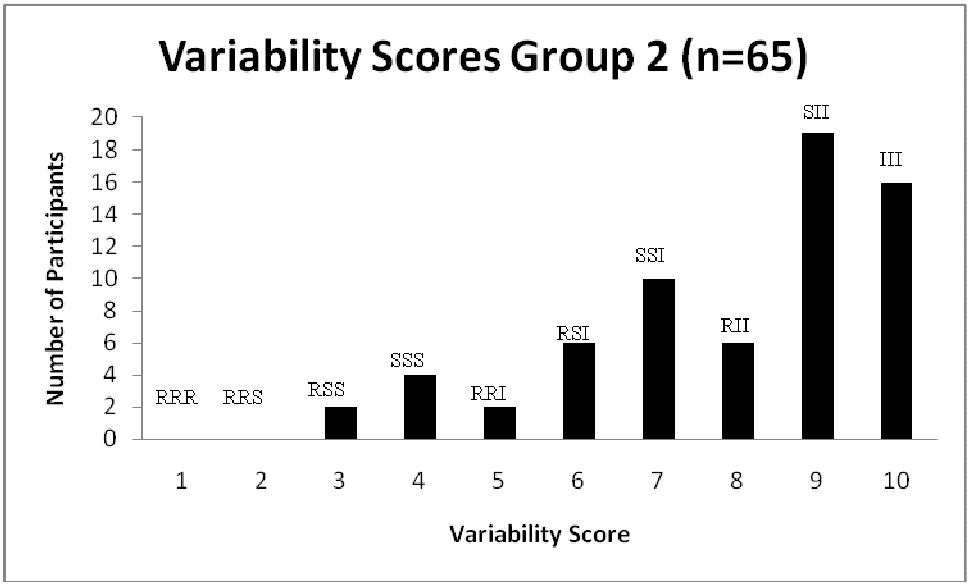


*Figure 2.2:* Number of participants in Group 2 achieving each variability score (1-10).

**Inferential Statistics**

**Between-Groups Differences.** A Mann-Whitney U Test revealed that Group 1 (Mdn = 5, n = 65) demonstrated significantly lower variability than Group 2 (Mdn = 9, n = 65), $U$= 895, $z$ = -5.72, $p$ = < 0.0005, $r$ = .5. This represents a large effect size (Cohen, 1988).

The data was then transformed according to the formula y = √10-x. The transformed scores were used to conduct an independent samples t-test. Levene's test for equality of variances was not significant so equal variance was assumed. The *t*-test revealed that Group 2 demonstrated significantly higher variability than Group 1 $t(128) = -6.183$, $p < .0005$, $r = .48$. This was very close to a large effect size.

**Correlations.** Kendall's tau was used to determine if variability scores from the Penny-Hiding Game were related to scores from the RBS-R for the participants with ASD (Group 1). There was a significant negative relationship between variability scores and total scores on the RBS-R, $\tau = -.24$, $p < .05$. Significant negative correlations were also observed between variability scores and scores on the Stereotypy Subscale, $\tau = -.28$ and the Self Injurious Behaviour Subscale, $\tau = -.28$ (all *p*s < .01). See Table 2 for significant correlations.

Variability scores were not significantly correlated with scores on the Compulsion Subscale, $\tau = .05$, the Ritualistic/Sameness Behaviours Subscale, $\tau = -.10$ or the Restricted Interests Subscale, $\tau = -.11$ (all *p*s >.05).

Finally, there was no significant correlation between the age of participants and their variability scores, $\tau = .001$, $p > .05$ across the entire sample.

| | **Variability Score** | **RBS-R Total Score** | **RBS-R Stereotypy score** | **RBS-R SIB Score** |
|---|---|---|---|---|
| Variability Score | 1.000 | -.243 ** | -.281 ** | -.281 ** |
| RBS-R Total Score | - | 1.000 | .510 ** | .486 ** |
| RBS-R Stereotypy Score | - | - | 1.000 | .369 ** |
| RBS-R SIB Score | - | - | - | 1.000 |

** $p < .01$ (2-tailed).

*Table 2.2:* Kendall's tau correlations between variability and scores on the RBS-R.

## 2.4. Discussion

The data reported above corroborate the findings of Miller and Neuringer (2000) and Baron-Cohen (1992) as the neurotypical group demonstrated significantly more variability on the Penny-Hiding Game than the ASD group. The results were robust with regard to statistical methods used, *t*-test or Mann-Whitney U Test. The larger sample size of the current study, together with the large effect size obtained, add significant support to previous research in this area.

The significant negative correlation between total scores on the RBS-R and variability scores from the Penny-Hiding Game reveal that high levels of RRBs are related to low variability among individuals with ASD. The analysis of relationships between variability scores and RBS-R subtypes revealed that low variability was related to higher levels of stereotypy and self-injurious behaviour. However, there was no significant relationship found between variability scores and compulsion, ritualistic/sameness behaviour or restricted interests.

These findings are interesting as it is clear that low variability was strongly related to the category of RRBs described by Lam et al. (2008) as RSMB. These are considered lower order RRBs. The other categories, IS and CI correspond to the four

subscales that were not significantly related to variability scores among this sample. As these categories are considered higher order RRBs it is possible that the participants who emitted more of this sub-type had the ability to vary their behaviour when it was functional for them to do so, i.e., when they wanted to “win” the game.

While some may believe that children with ASD are unlikely to be motivated by the social nature of this game, the experimenter observed throughout, with very few exceptions, that the children enjoyed participating in the game. They expressed satisfaction when they “won” and dismay when they “lost”. This game appealed to children across the spectrum of functioning, with many teachers reporting retrospectively that the children had requested the game or tried to play it with them following this study. Therefore, it may be concluded that a lack of motivation was not related to performance on the Penny-Hiding Game.

As RRBs have been reported to be in categories of lower and higher order, the findings of the current investigation may have important implications. The finding that lower order RRBs were related to lower variability warrants further consideration. Lam et al. (2008) noted that lower order RRBs were correlated with lower IQ in individuals with ASD. Future research should investigate the relationship between verbal and non-verbal IQ and variability as this may lead to important findings about the nature of low variability in ASD and ultimately lead to effective methodologies for increasing variability in different presentations of RRBs in this population. The next chapter will further investigate response variability in individuals with ASD using robust measurement techniques. Whilst the penny hiding game represents a naturalistic means of testing response variability in children, a computerised test along with an automated system of data collection should allow further analysis of response patterns. Furthermore, such a programme

can incorporate a statistical analysis using the U-value statistic that will yield an analysis of repetition and random responding. The U-value is a statistic that measures uncertainty or randomness in a data set. It gives a score from 0 to 1 with 0 being highly repetitive and 1 being highly variable (Neuringer, 2002). It allows for careful comparison between participants and this is desirable in order to expand the findings of the current study by utilising parametric statistics to detect between group effects. Other specific questions that are related to the findings of the current study will be addressed, including the effect of age on variability. Previous research suggests a relationship but no effect was observed in the current study. The next chapter will investigate this further using robust statistical methods and an automated data collection system.

# Chapter 3

## Study 2

## Investigating the Differences in Variability between Children with Autism Spectrum Disorder and Neurotypical Children: Relationships to Age and Receptive Language Ability

### 3.1. Introduction

As noted in the previous chapter, the presence of restricted repetitive behaviours (RRBs) is one of the diagnostic criteria for autism spectrum disorder (DSM IV-TR, American Psychological Association, 2000). Individuals who demonstrate high rates of RRBs can be said to have low variability in their behaviour as highly stereotypic responses are, by definition, invariable (Neuringer, 2002). Individuals with autism spectrum disorder (ASD) have been shown to have low variability in multiple repertoires, including verbal behaviour (Lee, McComas & Jawor, 2002) and play (Stahmer & Schreibman, 1992).

Numerous studies have investigated the impact of low variability among this population. Boucher (1977) asked neurotypical children (control group) and children with ASD (experimental group) to choose between two arms of a maze. The children in the experimental group tended to choose only one arm of the maze throughout, while the control children varied between the two arms. When a third arm was added, the children in the control group were far more likely to sample it than the experimental group, who tended to ignore the new option. This study demonstrated that children with ASD showed low variability in novel exploratory behaviour. Impairments in sampling and exploration may result in an inability to maximise the available reinforcement in different environments. Pierce and Courchesne (2001) described the narrow or restricted range of interests identified in children with ASD as a reduced tendency to explore and examine novel environments or elements of such environments. This could then lead to missed opportunities to learn from stimuli outside of a narrow range of interest.

Mullins and Rincover (1985) assessed the ability of children with ASD to sample available stimuli and maximise reinforcement in their immediate

environment. They asked participants with ASD and neurotypical controls to choose one of five cards. Each card was a discriminative stimulus for a different schedule of reinforcement (CRF, FR2, FR4, FR7, FR11). The control participants sampled all of the five cards and quickly learned to choose the schedule that produced the greatest reinforcement. The experimental group sampled fewer cards and often perseverated on a less optimal schedule. This study indicates the potential difficulties that result from low variability in sampling behaviour, most notably the tendency to "miss out" on a range of naturally occurring contingencies.

Furthermore, children with low variability in their sampling and exploratory behaviours may have difficulty in discovering response patterns that yield the greatest amount of reinforcement. In this way, they may fail to contact the contingencies for shaping such behaviour patterns and they may not develop appropriately, leading to disadvantages in this area into adulthood. For example, if a child with ASD does not sample and explore their environment, they may miss out on the shaping processes that result in neurotypical children developing such exploratory behaviour (Skinner 1938). In this way, they may fail to encounter opportunities for greater access to reinforcers and the development of novel interests and experiences. This may have an adverse effect across the lifespan with the resultant narrow range of circumscribed interests affecting their social integration. Neuringer (2002) described the inability of the children in the Mullins and Rincover (1985) study to emit the type of behaviour that would locate the optimal schedule of reinforcement as "a failure to adapt" (p. 687). This may have implications for a person's ability to be successful in social, academic and occupational pursuits.

The various types of RRBs emitted by an individual diagnosed with ASD are thought to be related to various aspects of their clinical presentation. Turner (1999)

referred to stereotypical behaviours, object manipulations and self-injurious behaviours as lower order behaviours that are related to more significant impairment in intellectual functioning. She described traits such as restricted interests and unusual attachments to objects or toys as higher order RRBs that are related to higher intellectual functioning (see also Lam, Bodfish & Priven, 2008). Gabriels, Cuccaro, Hill, Ivers and Goldson (2005) reported that non-verbal IQ was negatively correlated with the Sameness Behaviour subscale of the Repetitive Behaviour Scale - Revised (Bodfish, Symons & Lewis, 1999). Lam (2004) reported a lack of published literature examining the effect of cognitive functioning on the presence or severity of repetitive behaviours. This is an area that warrants further investigation in order to improve our understanding of the factors that influence variability in behaviour, across individuals with autism, and indeed other clinical presentations.

Miller and Neuringer (2000) directly reinforced variability in responding on a computer game and compared the performance of adolescents with ASD to that of adult and child controls. They found that participants with ASD responded significantly less variably overall than the adult control group and, while they showed less variability than the child control group, this did not reach statistical significance. However, this study did not evaluate participants' performance in relation to differential participant characteristics.

Analysis of the literature pertaining to operant variability reveals a dearth of investigations on the effects of variables such as age and gender on variability in responding. Sutton-Smith (1975) described early play behaviours as repetitive and reported that play became more complex and varied as the child grew older. In 1960, Hanlon compared groups of children at two different ages (3-4 years and 6-7 years) using a checkerboard maze. She found that the older children showed greater

variability in the paths they chose across the board. More recently, Bancroft (2011) investigated variability in selection responses on a variety of tasks including selecting different coloured crayons for colouring, selecting different outfits for a dressing task and selecting beads for threading in children with autism and neurotypical children. Outcomes revealed that there was no significant difference in variability as a function of gender. The author noted that in the case of the neurotypical children, as age increased, so did variability in responding, i.e., older participants were more likely to be classed as "high variability". However, this outcome was not observed in the sample of children with ASD as variability remained stable with age in this group.

While there remains a paucity of operant variability research with human participants, the basic research has considered the question of age as a moderating variable on response variability. Neuringer and Huntley (1992) examined the effects of age, gender and contingency on the variability of sequences of lever presses in rats. They found that younger rats showed higher variability in response patterns and that there was no gender difference in response variability. Variability was sensitive to contingencies with the rats demonstrating higher levels of variability during the VAR condition, where reinforcement was contingent on variation, than in the YOKE condition, where reinforcement was independent of variation. The finding that younger rats showed higher variability is contrary to the human research described above but further research on this phenomenon is warranted.

The current study aimed to further investigate the difference in variability between children with ASD and neurotypical children. Specifically, in order to more carefully control the conditions under which variability was measured, a computer program in the form of a game, adapted from Miller and Neuringer (2000), was

developed for this purpose. In this study, the program calculated a U-value statistic as a measure of variability in responding for each participant in order to determine the distribution of response probabilities.

The U-value examines this distribution, with equal probabilities indicating high variability and unequal probabilities indicating low variability (Neuringer, Kornell & Olufs, 2001). In effect, the U-value approaches 1 when frequencies are approximately equal, and approaches 0 when one possibility predominates (Neuringer, 2004). In other words, higher U-values (closer to 1) indicate higher variability while lower U-values (closer to 0) indicate lower variability or more repetitive responding. By comparing U-values, the current study aimed to examine the difference in variability between the experimental and control groups. Additional analysis was conducted, based on supplemental measures, in order to investigate whether low or high levels of variability were related to language ability, age or sex in children with ASD and separately in children of typical development.

## 3.2. Method

### Participants

This study involved three participant groups with a total of 30 participants. The experimental group (EXP) comprised 10 children with a diagnosis of ASD (6 males and 4 females) aged between 5 years, 9 months (69 months) and 13 years, 4 months (160 months) with a mean age of 106 months (SD = 8.5 months). All of the participants in the experimental group had been assessed by an independent psychologist and had been found to meet the DSM-IV diagnostic criteria for ASD. All participants with ASD attended an ABA school for 30 hours per week and had received between two and five years of ABA intervention at the time of the study.

The participants in this group all had impaired verbal behaviour relative to their chronological age. They communicated via picture exchange or vocal utterances.

The first control group was matched to the EXP group for sex and chronological age (CCA). It comprised 10 neurotypical children who attended a mainstream primary school without any additional supports. They ranged in age from 9 years, 2 months (110 months) to 13 years, 5 months (161 months) with a mean age of 143 months (SD = 6 months). The second control group (CAE) was matched for sex and age equivalence as indicated by the age equivalent scores obtained by the EXP group on the Peabody Picture Vocabulary Text – 4 (PPVT-4). This group comprised 10 neurotypical children, all of whom attended either a mainstream preschool or mainstream primary school without any additional supports. They ranged in age from 3 (36 months) to 6 years, 7 months (79 months) with a mean age of 71.5 months (SD = 3.5 months). This group was included to control for the effect of language ability as many of the participants in the EXP group had delayed verbal repertoires. The matching procedure utilised in this study resulted in matched triads whereby each participant in the EXP group had two controls, one in the CCA group and one in the CAE group. Table 3.1 lists the set of 10 triads.

| Experimental Group (EXP) | | | Controls Matched for Chronological Age (CCA) | | | Controls Matched for Age Equivalence (CAE) | | |
|---|---|---|---|---|---|---|---|---|
| **Participant number** | **Age in years and months** | **Sex** | **Participant Number** | **Age in years and months** | **Sex** | **Participant number** | **Age in years and months** | **Sex** |
| **1** | 12:2 | F | **11** | 12:5 | F | **21** | 6:3 | F |
| **2** | 13:4 | F | **12** | 13:5 | F | **22** | 4:3 | F |
| **3** | 6:5 | F | **13** | 6:2 | F | **23** | 6:7 | F |
| **4** | 5:9 | F | **14** | 5:6 | F | **24** | 3:1 | F |
| **5** | 8:10 | M | **15** | 9:3 | M | **25** | 3:1 | M |
| **6** | 6:10 | M | **16** | 7:2 | M | **26** | 3:1 | M |
| **7** | 6:7 | M | **17** | 6:4 | M | **27** | 6:3 | M |
| **8** | 7:0 | M | **18** | 7:10 | M | **28** | 3:0 | M |
| **9** | 11:8 | M | **19** | 11:11 | M | **29** | 6:0 | M |
| **10** | 11:2 | M | **20** | 11:9 | M | **30** | 4:9 | M |

*Table 3.1:* List of matched triads along with demographic information (age and sex).

All 30 participants had the ability to sit at a table appropriately and attend to an activity or person for extended periods of time (>15 minutes).

## Measures

**The Peabody Picture Vocabulary Test Fourth Edition Form A (PPVT-4; Dunn & Dunn, 2007).** The PPVT-4 is a norm-referenced instrument containing 228 test items that measure the receptive vocabulary of children and adults. Participants are required to discriminate one target from an array of four pictures by either pointing to the picture or by saying the number corresponding to their picture of choice. The PPVT-4 does not require the participant to interact verbally with the examiner and reading and writing are not necessary. It is a useful tool for participants who have difficulties with these particular skills (Dunn & Dunn, 2007).

The PPVT-4 yields a Raw Score which can be converted into a Standard Score and, in turn, into an Age Equivalent Score.

A norm sample of 3,540 individuals was collected to test the PPVT-4. They were aged between 2 years, 6 months and 81 years. From aged 2 years, 6 months to 61 years a target of 50% male and 50% female was used and from ages 61 to 81 years target percentages were based on the US population. The PPVT-4 has been shown to have good internal consistency. The split-half reliability was calculated for Form A and Form B of this test across the entire normative sample. These were shown to be very consistent across both forms and averaged .94 or .95, which is considered very high. The authors note that reliabilities for the preschool and kindergarten age groups were “as high, if not higher” than for the older age groups.

In earlier versions of the instrument, the age equivalent score was referred to as “mental age”. This instrument was also used in many published studies as a measure of non-verbal IQ. For example, Reed, Broomfield, McHugh, McCausland & Leader (2009) used the British Picture Vocabulary Scale (BPVS, Dunn, Dunn, Whetton & Pintilie, 1982), which is derived from the PPVT but standardised for use with children in the UK, as a measure of mental age and the authors also made comments pertaining to the IQ of their participants in relation to their scores on the BPVS. In the newest edition, the authors no longer refer to the score as an IQ score but a measure of receptive language. The PPVT-4 was chosen for the current purpose due to its ease of administration and its suitability for the EXP group, many of whom had limited verbal repertoires.

## Apparatus and Materials

**Hardware for program presentation.** A Hewlett-Packard laptop (HP Pavillion dv6000) using Windows Vista operating system was used in this study. It had 512. 0 MB RAM with processing speed 1600. 0 MHz. The screen measured 33cm x 22cm.

**Response hardware (input device).** An Ergodex DX1 input system with programmable response keys was placed 2 cm in front of the laptop. It had two buttons (L & R) placed 8 cm apart.

**Computer program.** A program in the form of a game to measure variability in responding was loaded onto the computer. This program was designed to the specifications of the experimenter for the purpose of the current study. It recorded every sequence emitted by the participants and used this data to calculate the U-value statistic for each individual participant according to the following formula:

$$U = -\sum[(N_{1\text{-}16})*(\log_2[N_{1\text{-}16}])]/\log_2(16)$$

The developer was a qualified technician and an employee of the National University of Ireland Galway. He had prior experience and expertise in developing a range of data driven applications for research purposes. The developing environment was Visual Studio 2008 (Express edition) and the programming language used was VB.NET (.NET version 3.5).

## Setting

The participants in the EXP group completed all components of the study in a quite area of their classroom that was separated from the main classroom by a screen and only the student and experimenter were in this area. The control children, who

attended primary school, completed all components of the study in the office of the school secretary. Only the student and the experimenter were present with the door open and the secretary working in the hallway outside. The children who attended preschool completed all components in a playroom attached to their classroom. Only the student and the experimenter were in the room but there was a large window so that the teachers could observe the scene at all times. Participants sat facing the computer and the experimenter sat approximately 1 metre away.

## Procedure

The Peabody Picture Vocabulary Test (PPVT) was administered to the ten participants in the experimental group. This measure yielded an age equivalence score that was used to select participants for the group matched for age equivalence (CAE). Ten neurotypical children were matched for sex and for age-equivalence in this way. A further ten neurotypical children were matched for sex and chronological age. The PPVT was then administered to the participants from each control group to obtain standard scores for later statistical analysis.

Following the PPVT, all participants completed the computer program. The neurotypical children completed the PPVT-4 and the computer program on the same day, while the children with ASD completed them within one week of each other.

**Computer Program.** Before the session commenced the following instructions appeared on the computer screen:

*"In this game you must make happy faces appear on the screen by pressing the two buttons. If you can fill the big circle on the screen with faces then you will get a prize. If you see a red box on the screen you must wait a second before pressing the buttons"*

The experimenter pressed one of the buttons to start the game and a large circle appeared in the centre of the screen. It contained the outline of 15 smaller circles in which a smiley face appeared when a trial was reinforced. In this study, one trial consisted of a sequence of four presses on either left (L) or right (R) buttons e. g., LLLR, RLRL, RRLL. There were 16 possible sequences of four presses. Each button emitted a different tone of .2 seconds duration when pressed. Table 3.2 lists all 16 possible sequences of four presses across the two buttons.

| **Sequence Options** |
|---|
| 1. LLLL |
| 2. LLLR |
| 3. LLRR |
| 4. LRRR |
| 5. RRRR |
| 6. RRRL |
| 7. RRLL |
| 8. RLLL |
| 9. RLRL |
| 10. LRLR |
| 11. LRRL |
| 12. RLLR |
| 13. LRLL |
| 14. RLRR |
| 15. LLRL |
| 16. RRLR |

*Table 3.2:* List of all 16 possible sequences of four presses using two buttons.

After each sequence of four presses there was a pause of one second, during which a red box (6cm x 6cm) with STOP written within it appeared on the screen. During this time, if the participant pressed either button no tone sounded and this

press was not counted as part of the next trial. If a button was pressed while the STOP box was on the screen the box remained there until one second had passed without any buttons being pressed.

Each time a trial was reinforced a yellow “happy face” appeared in one of the small circles (see Figure 3.1). The criterion for finishing the computer program was to obtain all 15 “happy faces”. At this point music played, balloons floated around the screen and the happy faces became animated, smiling and winking. This lasted for 30 seconds and then an opportunity was presented for the students to select a secondary reinforcer from their choice board. The choice boards used for the EXP group were the same ones they used during their school day. These were maintained as the items on the choice board had been identified through a preference assessment. The control participants were also given an opportunity to select a reinforcer from a choice board. These choice boards were constructed following consultation with the appropriate teachers. They contained options for selecting preferred toys or books to engage with for two minutes. There were two settings for the computer program – training and testing.

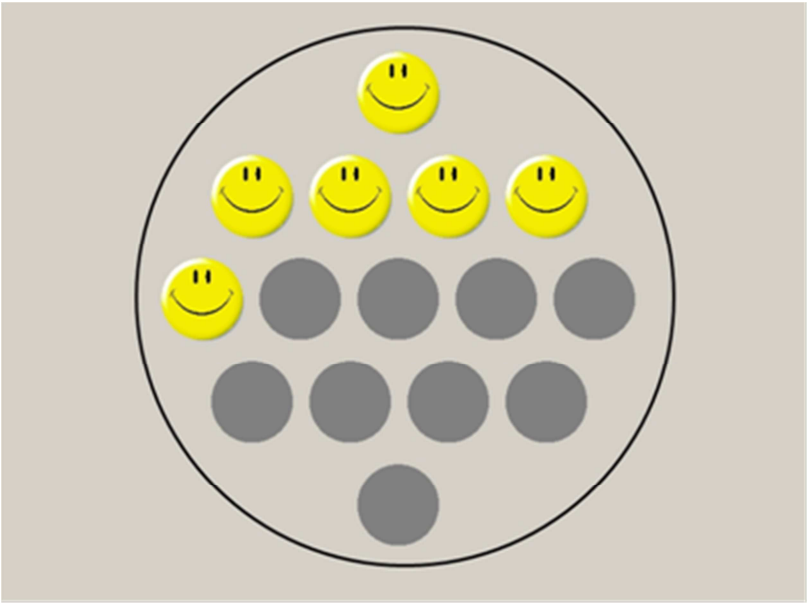

*Figure 3.1:* Screen shot of computer program during trials.

***Training Phase.*** A training phase was conducted to teach the participants how to use the computer program, in particular to teach them to respond in sequences of four presses. During the training phase, the experimenter read through the instructions with the participant. If the participant was only responding on one button during the first game the experimenter said "remember there are two buttons" but did not intervene in any other way unless the participant stopped responding, in which case the experimenter said "let's keep going to see what happens at the end and get your prize". This phase consisted of 15 trials per game with an FR1 schedule of reinforcement in operation. After 15 trials the criterion was reached and the 30 second music and animation was accessed. A tangible reinforcer was also delivered at this point. This served as a conditioned reinforcer for engaging with and completing the program.

In the training section a correct trial was defined as a sequence of 4 button presses followed by a one second pause while the STOP box was on the screen. The experimenter observed and recorded the frequency of correct trials during the training phase. The criterion for progressing to the testing phase was set at 10 correct trials out of 15 during the training phase. All participants reached this criterion in two training sessions or less.

***Testing Phase.*** In this phase the computer game consisted of 30 trials with 50% of trials obtaining reinforcement. In each game, 14 reinforcers in the form of "happy faces" were delivered randomly across the first 29 trials. The 30$^{th}$ trial was always reinforced to complete the circle (15 "happy faces") and signify the end of the game. At this point the music and animation occurred and a tangible reinforcer was delivered as in the training session. The differences between the training and

testing phases were that there were 15 trials to complete a game in the training phase and 30 trials in the testing phase. In the training phase every trial was reinforced while in the testing phase 50% of trials were reinforced with reinforcement randomly distributed.

Each participant completed five games in the testing phase, resulting in 150 trials for each participant. The computer program provided the following output for each participant:

- a summary of the sequence emitted for every trial;
- whether each sequence was reinforced or not;
- a summary of each of the 16 possible sequences and the frequency at which each one occurred;
- a U-value for each set of 150 trials.

**Data Analysis**

A within group ANOVA was conducted to examine the differences in variability between the experimental, chronological age control group and age equivalence control group. Further statistical analysis was conducted to determine if higher variability scores were correlated with a higher standard score on the PPVT-4. Pearson product-moment correlations were conducted to investigate if there were relationships between U-value and non-verbal IQ and U-value and age among the entire group (n=30). Furthermore, correlations were conducted to determine if there were relationships between these factors among the neurotypical group (CCA + CAE, n=20) and among the EXP group (n=10) separately.

### 3.3. Results

Figure 3.2 displays the mean age in months for each group. Included in the graph is the mean age equivalent score for the EXP group, demonstrating that, as a group, they performed significantly below their chronological age on the PPVT-4. Thus, the mean age of the control group matched for age equivalence (CAE) is significantly lower than that of the control group matched for chronological age (CCA). This meant that a range of ages was represented in the sample of neurotypical controls (one group with a mean age of 9 years and one group with a mean age of 4 years).

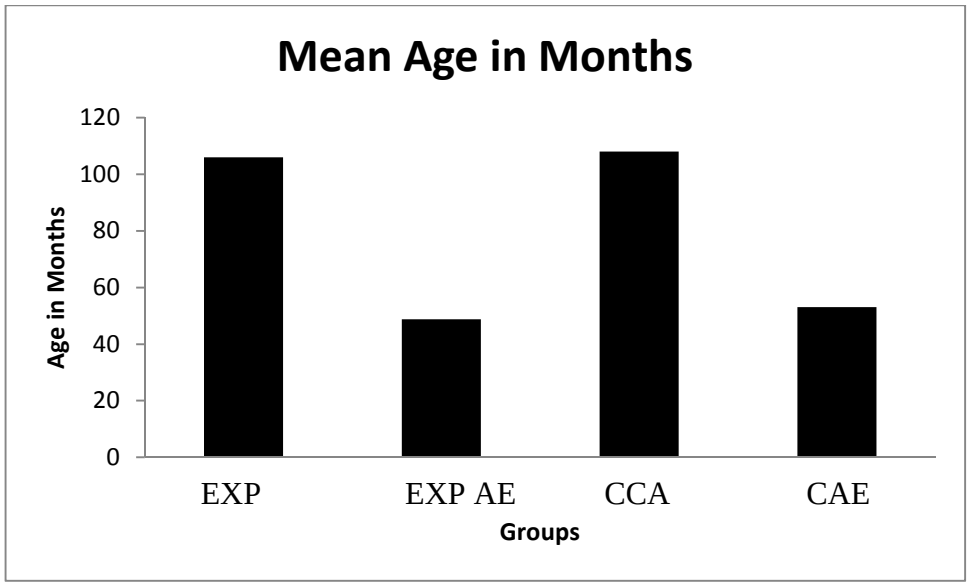


*Figure 3.2.* Mean age in months of EXP (experimental group), EXP AE (age equivalence scores of the EXP group), CCA (chronological age matched controls) and CAE (age equivalence matched controls).

Table 3.3 summarises the mean U-value, along with the standard deviations, for each group. The mean U-value for each group is represented in Figure 3.3. For the EXP group the mean U-value was 0.45, SD = 0.24 (range: 0.07 – 0.80). The

mean U-value for the CCA group was 0.59, SD = 0.29 (range: 0.08 – 0.93), while the mean U-value for the CAE group was 0.73, SD = 0.19 (range: 0.30 – 0.94).

| **Group** | **Experimental (EXP)** | | **Control – Chron. Age (CCA)** | | **Control – Age Equiv. (CAE)** | |
|---|---|---|---|---|---|---|
| | Mean | Standard Deviation | Mean | Standard Deviation | Mean | Standard Deviation |
| **U-value** | 0.45 | 0.24 | 0.59 | 0.29 | 0.73 | 0.19 |
| **Standard Score** | 53.10 | 24.01 | 103.40 | 13.31 | 108.70 | 11.98 |
| **Age in months** | 105.90 | 35.15 | 108.00 | 36.91 | 53.00 | 16.57 |

*Table 3.3:* Mean U-values and Standard Deviations for the three groups.

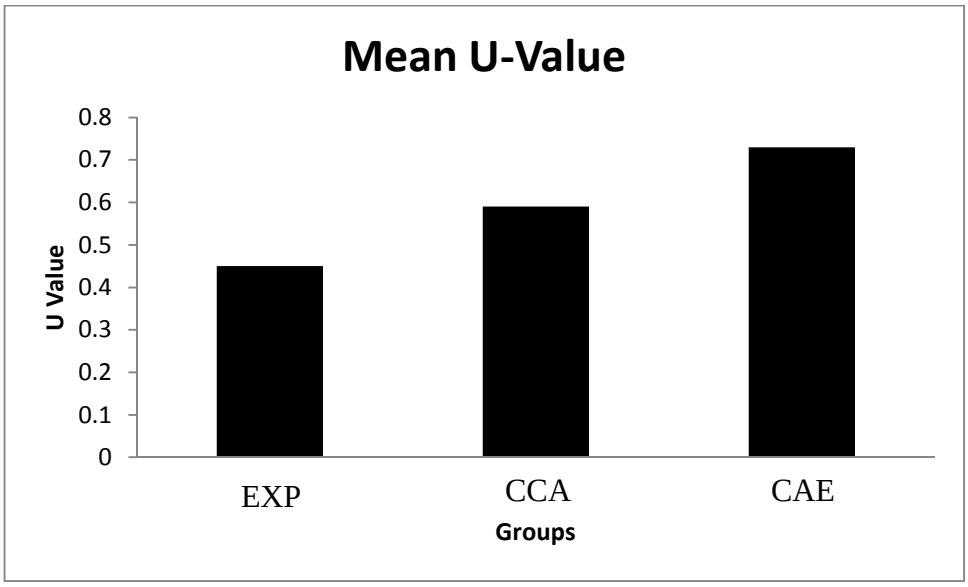


*Figure 3.3:* Mean U-value for each group (Experimental, Chronological Age controls and Age Equivalence controls).

Figure 3.4 represents the mean standard score on the PPVT-4 for each group. The mean standard score for the EXP group was 53.10, SD = 24.01 (range: 30 – 103). The mean standard score for the CCA group was 103.40, SD = 13.31 (range:

82 - 123). Finally, the mean standard score for the CAE group was 108.70, SD = 11.98 (range: 91 – 128).

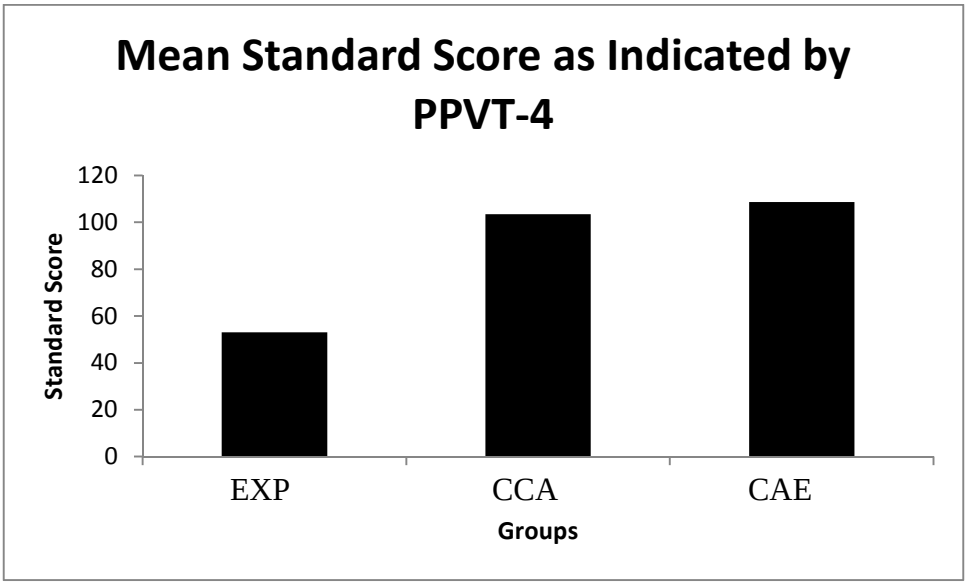


*Figure 3.4*: Mean standard score (SS) for the 3 groups (Experimental, Chronological age controls and Age equivalence controls.

A one-way ANOVA revealed a significant difference in U-value between groups ($F(2,27) = 3.\ 53$, $p = .043$). A Tukey post hoc test revealed that the U-value was statistically significantly higher in the control group matched for mental age (.75 ± .19, p = .034) than in the experimental group (.45 ± .24). There were no statistically significant differences in U-value between the experimental group and the control group matched for chronological age (.59 ± .28), or between the two control groups.

A Pearson product-moment correlation was conducted to determine if there was a relationship between participants' U-value and their language ability as indicated by standard score on the Peabody PVT-4. There was no correlation

between U-value and standard score ($r$ = .251, $p$ = .181). A second Pearson product-moment correlation was conducted to determine if there was a relationship between the U-value and age among the entire participant group (n = 30). There was a significant negative correlation between U-value and age ($r$ = -.402, $p$ = .027).

Data from the two control groups combined (n = 20) were analysed to determine whether standard score on the PPVT-4 or age would be related to the U-value in a neurotypical population. A Pearson product-moment correlation was conducted and no significant relationship was detected between U-value and standard score ($r$ = -.065, $p$ = .785) or between U-value and age ($r$ = .203, $p$ = .391) among this group.

Finally, the data from the experimental group was analysed to investigate the same relationships in a population with autism. A Pearson product-moment correlation was conducted and there were no significant correlations between U-value and standard score ($r$ = -.272, $p$ = .447) or between U-value and age ($r$ = -.618, $p$ = .057) in this sample.

### 3.4. Discussion

The results of Study 3 are surprising for a number of reasons. Firstly, it was expected that the experimental group would demonstrate significantly lower variability than both of the control groups. While a significant result was revealed by the analysis of variance (ANOVA), the post hoc test showed that there was a statistically significant difference in U-value between the experimental group and the age equivalent control group but none existed between the experimental group and the chronological age controls. Furthermore, there was a significant negative correlation between U-value and age indicating that the younger participants

demonstrated higher variability. This is contrary to previous reports that variability increases with age (Bancroft, 2011). However, when the data from the experimental group and the combined control groups were analysed separately, no such significant effect for age was observed for either set of participants.

It was hypothesised that the standard score on the PPVT-4 would be positively associated with U-value. This association was not observed in the current study. This indicates that higher receptive language ability (formally referred to as "non-verbal IQ") was not found to be associated with increased behavioural variability in this study. Further investigation is warranted to answer this research question should researchers be interested in identifying factors associated with low or high variability.

Interestingly, results revealed that although the control group matched for age equivalence (CAE) were matched to the EXP group according to their age equivalent scores on the PPVT-4, the CAE group demonstrated significantly higher standard scores than the EXP group on the PPVT-4, indicating higher language ability. This indicates that the CAE group performed better than expected for their age on this measure. This discrepancy must be considered when interpreting the outcomes of the study.

The outcomes of this study do not fully support the findings of Miller and Neuringer (2000), whose data suggested that children and adults of typical development were significantly more variable than adolescents with ASD. In the current study, only the CAE group were significantly more variable than the EXP group. This indicates that the control group, who had a lower mean age, differed significantly from the participants with ASD although the control group matched for chronological age (CCA), who had a higher mean age, did not. However, visual

inspection of the data reveal that the participants with ASD did indeed demonstrate lower variability than both control groups but the effect did not reach statistical significance for the CCA group. The aim of this study was to investigate whether children with ASD would demonstrate lower variability than their neurotypical same-age peers and age equivalent peers. The outcomes indicate that this is the case.

A measure of age equivalence determined by receptive language ability using the PPVT-4 may have yielded a somewhat limited language ability score. Nevertheless, the ease of administration of the PPVT-4, especially with the ASD group, meant that it was not problematic to include it as a variable in the current analysis. As there were no significant effects we can conclude that there was no effect of language ability on variability in this sample. However, the extent to which we can generalise this finding to cognitive functioning is limited and could be the subject of further analysis in the future.

It became clear in the course of the study that the design of the computer program may have affected the motivation of the older participants. The majority of the participants in the control groups reported playing computer games in their spare time. The games that the older participants played are sophisticated and complex and the current program may have appeared partially rudimentary in comparison. The secondary reinforcer was added to address this issue and, anecdotally, the experimenter can report that the participants willingly engaged with the program throughout the study.

**Future directions**

The current study raises a number of areas for further investigation. Firstly, the effects of age and IQ on variability merit additional investigation across neurotypical populations and also with individuals with ASD. Considerably larger sample sizes would be desirable. Indeed a large scale study or a longitudinal one, tracking changes in age, IQ and variability over time, is likely to address some of the inconsistencies in findings that exist amongst some of the studies utilising smaller sample sizes, including the current one. Further knowledge of the developmental trajectory of variability within an ASD population will yield more information on the moderating variables that affect it over time.

Considering the appeal of technology to children and adolescents with and without an ASD diagnosis, this is an area that warrants further investigation. It is likely that the development of appealing computer programs for use with the clinical population would lead to a high level of motivation for engaging with such programs. In this way individuals could practice response variation at the level of the computer program at least and this may increase their ability to vary across other repertoires of behaviour.

Finally, additional research that investigates the effects of low variability on the adaptive functioning of individuals with ASD would add to the literature in a positive way. It is important for researchers and clinicians to understand the impact of low variability for individuals with ASD and their families. The development of methods for ameliorating low variability in individuals with ASD may have important implications for their overall learning and development. The next two chapters will investigate methods for achieving this aim. In Chapter 4, a sophisticated computer program incorporating lag schedules of reinforcement will be

utilised. In Chapter 5, two experiments will be described that apply the findings of the previous chapters to an applied setting and a socially significant behaviour, variability in play actions with toys. Chapter 4 will address some of the limitations of the current chapter by incorporating a professional animator and designing an appealing and engaging interface. Furthermore, the effects of different lag schedules on the variability in responding of participants with a diagnosis of ASD and their neurotypical peers will be evaluated in order to identify potential remediation strategies for the low variability commonly observed in individuals with ASD.

# Chapter 4

# Study 3

# Increasing Variability in Children with Autism Spectrum Disorder and Neurotypical Children Using Lag Schedules of Reinforcement

## 4.1. Introduction

Lag schedules of reinforcement have been used to increase variability in responding in both basic (Neuringer & Huntley, 1992; Neuringer, 1992; Cherot, Jones & Neuringer, 1996) and applied (Cammileri & Hanley, 2005; Napolitano Smith, Zarcone, Doodkin & McAdam, 2010) research in the field of behaviour analysis. Lag schedules involve the delivery of reinforcement if a response, or sequence of responses, differs from a pre-determined number of responses (Page & Neuringer, 1985). For example, under a lag 3 schedule, a response would be reinforced if it differed from the previous three responses while under a lag 10 schedule, a response would be reinforced if it differed from the previous 10 responses.

Lag schedules have also been used to investigate the operant nature of variability, commonly compared with a control condition where reinforcement is provided independent of variability. To date the majority of the publications incorporating lag schedules are in the area of basic research. Increasingly, these data have begun to inform the applied research, with several studies emerging in recent years (for example, Cammileri & Hanley, 2005; Lee & Sturmey, 2006; Napolitano et al., 2010) detailing the use of lag schedules to increase variability in play and language in the clinical population. The progression in utilising such procedures to increase variability in response patterns, from basic to applied research, is indicative of the value that researchers and practitioners are now placing on variability as a target behaviour in humans. Furthermore, lag schedules may offer an efficient and reliable tool for increasing response variability when this type of responding is necessary or desirable.

While several studies have emerged that employ lag schedules to increase variability, the majority of the applied research has been limited to a lag 1 schedule. In light of the outcomes that basic researchers have demonstrated (for example, inducing marked increases in response variability with pigeons using lag 5 and lag 10 schedules (Abreu-Rodrigues, Lattal, Dos Santos & Matos, 2005) and even a lag 50 schedule (Page & Neuringer, 1985)), the use of the higher lag values with human participants is an obvious area for further investigation. In 2012, two research papers were published involving the use of lag schedules to increase variability in the verbal behaviour of participants with developmental disabilities. Susa & Schlinger (2012) used lag schedules of reinforcement to increase the intraverbal responses of a boy with autism spectrum disorder (ASD). They increased the lag requirement from lag 1 to lag 2 and then to lag 3 in a changing criterion design. They observed that the number of previous responses from which the current response differed increased steadily as the lag requirement was increased. Heldt and Schlinger (2012) increased variability in the tacting of two boys, one of whom had mild intellectual disability, the other of whom had Fragile X syndrome. They implemented a lag 3 schedule directly after the baseline phase and they noted increased variability in tacting for both participants. Furthermore this increase was maintained at a three-week follow-up. These studies illustrate the beginning of a move toward the increased use of these types of strategies in the applied field.

The outcomes reported in the previous two chapters indicate that children and adolescents with ASD demonstrate deficits in variability in responding. Such findings are consistent with previous research and indicate a repertoire of behaviour that warrants intervention. Possible moderating factors identified include the age of the individual (although there is contrasting evidence as to the nature of this effect)

and the level and type of RRBs emitted. The current study aimed to identify possible remediation strategies for this effect. To this end, it investigated the effect of lag schedules of reinforcement on the variability of response sequences of children with ASD on a computer task. The task involved a computer program in the form of a game that required participants to fulfil a lag criterion in order to progress through it. A primary aim of the study was to determine if increasing the lag criterion would increase the level of variability as measured by the U-value statistic. It was expected that U-values would be higher when variability was contingent on increased variability, i.e., higher lag values would produce higher U-values. A secondary aim of the study was to compare how two groups of participants, one comprised of neurotypical children and one comprised of children with ASD, would perform on this task, with the neurotypical children expected to demonstrate higher variability than the children with ASD. Following on from the findings in Chapter 2 and Chapter 3, the differences in responsiveness across the two groups to altered lag schedules, was also of interest in this study.

## 4.2. Method

### Participants

Ten children with a diagnosis of ASD and ten neurotypical (NT) children participated in this study. The participants ranged in age from 5 years 2 months (62 months) to 15 years 1 month (181 months) at the time of the study (mean age = 107.5 months, SD = 43.8 months). The ASD group was labelled Group 1 while the NT group was labelled Group 2. In Group 1, four of the participants attended an Applied Behaviour Analysis School for children with ASD; three attended a special school for children with ASD that did not have ABA as its primary methodology and three

of the participants attended mainstream school and received some additional resource hours. All of the children received a diagnosis of ASD according to the DSM-IV criteria from a psychologist who was independent of this study. There were nine boys and one girl in Group 1. All of the participants in Group 1 were reported to have basic literacy skills.

The participants in Group 2 all attended a mainstream school. None were diagnosed with ASD or any other developmental or behavioural disorder. None of the participants in Group 2 were in receipt of additional supports in their school. The participants in this group ranged in age from 6 years, 1 month (73 months) to 12 years 1 month (145 months) at the time of the study (mean = 113.4 months, SD = 8.5 months). Group 2 was comprised of eight boys and two girls. All of the participants in Group 2 had basic literacy skills. Table 4.1 describes the participant characteristics of the two groups.

| **Participant Number** | **Group** | **Sex** | **Age (years:months)** | **ADOS-R Scores** | | |
|---|---|---|---|---|---|---|
| | | | | **Comm.** | **Social** | **Total** |
| **1.** | 1 | Female | 10:4 | 3 | 6 | 9 |
| **2.** | 1 | Male | 8:0 | 3 | 5 | 8 |
| **3.** | 1 | Male | 8:0 | 5 | 7 | 12 |
| **4.** | 1 | Male | 9.2 | 10 | 18 | 28 |
| **5.** | 1 | Male | 15:1 | 2 | 6 | 8 |
| **6.** | 1 | Male | 5:2 | 3 | 6 | 9 |
| **7.** | 1 | Male | 8:2 | 8 | 11 | 19 |
| **8.** | 1 | Male | 8:9 | 7 | 7 | 14 |
| **9.** | 1 | Male | 5:7 | 4 | 9 | 13 |
| **10.** | 1 | Male | 10:11 | 9 | 11 | 20 |
| **11.** | 2 | Male | 10:3 | n/a | n/a | n/a |
| **12.** | 2 | Male | 11:2 | n/a | n/a | n/a |
| **13.** | 2 | Male | 10:4 | n/a | n/a | n/a |
| **14.** | 2 | Male | 12:0 | n/a | n/a | n/a |
| **15.** | 2 | Female | 12:1 | n/a | n/a | n/a |
| **16.** | 2 | Male | 6:7 | n/a | n/a | n/a |
| **17.** | 2 | Male | 6:1 | n/a | n/a | n/a |
| **18.** | 2 | Female | 7:8 | n/a | n/a | n/a |
| **19.** | 2 | Male | 9:1 | n/a | n/a | n/a |
| **20.** | 2 | Male | 9:3 | n/a | n/a | n/a |

*Table 4.1:* Participant information for Group 1 (age, sex and ADOS-G scores) and Group 2 (age and sex)

## Apparatus

**Program Hardware.** A Hewlett-Packard laptop (HP Pavillion dv6000) using Windows Vista operating system was used in this study. It had 512.0 MB RAM with processing speed 1600.0 MHz. The screen measured 33cm x 22cm.

**Response hardware (input device).** An Ergodex DX1 input system with programmable response keys was placed two cm in front of the laptop. It had four keys placed in a row and each were 2cm apart.

**Computer program.** A program in the form of a game to measure variability in responding was loaded onto the computer. The program had six settings, five of which corresponded to a different value of the lag schedule. The sixth setting was a control condition where reinforcement was provided independent of variability. The program provided output in the form of a U-value statistic for each of the six settings.

***Development of the computer program.*** This program was designed to the specifications of the experimenter for the purpose of the current study. A professional animator was hired to improve the quality of the program, i.e., to reflect a typical computer game. The images were created using Adobe Illustrator. It was decided that animals would be used as the characters, as they are appropriate for a wide age range and they are immune from gender, racial and age biases that the participants may have had.

The colour palette was varied but consisted mainly of primary colours. The animator designed 10 “levels”. See Table 4.2 for a description of the levels. In each level, the character took 10 movements/steps to reach the target item and this was followed by a subsequent level appearing on the screen.

| Level | Description |
|---|---|
| 1. | Dog searching in holes in a garden for a bone. |
| 2. | Monkey swinging from tree to tree to reach a bunch of bananas. |
| 3. | Bear climbing up a ladder and across a gangway to reach a treehouse. |
| 4. | Bee flying from flower to flower to reach a jar of honey. |
| 5. | Frog leaping from stone to stone to across a pond to reach some insects. |
| 6. | Worm moving through an underground system of passages to find an apple. |
| 7. | Rocket flying from planet to planet to reach a birthday party. |
| 8. | Mouse running and climbing through the mouse-holes in a room to reach some cheese. |
| 9. | Octopus swimming around rocks and plants to find a friend. |
| 10. | Hen climbing up through a chicken coop to reach some eggs. |

*Table 4.2:* Description of each level of the computer program.

Once the drawings were completed by the animator, a developer created the program following consultation with the experimenter. As in Study 3, the developer was an employee of the University and had expertise in developing a range of data driven applications for research purposes. The developing environment was Visual Studio 2008 (Express edition) and the programming language used was VB.NET

(.NET version 3.5). Figure 4.1 illustrates the apparatus as it was presented to the participants.

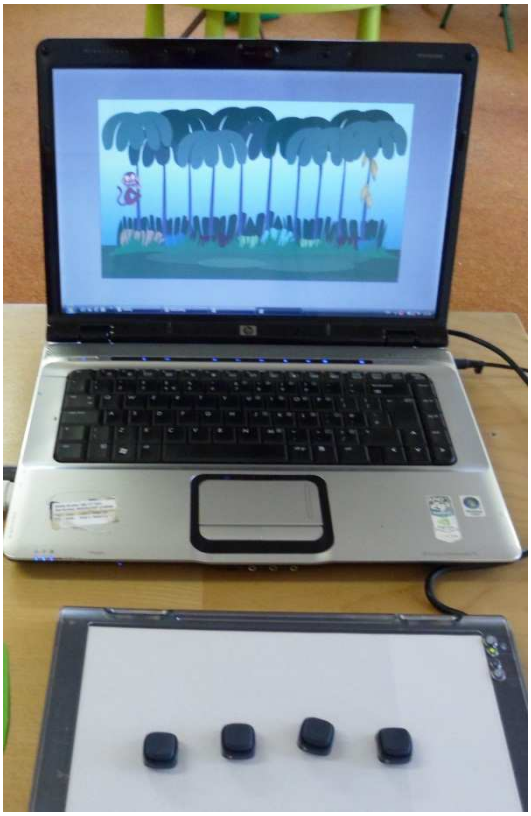

*Figure 4.1:* Apparatus: The laptop and Ergodex® with the program on display.

An important consideration in designing the program was the number of trials (sequences of key presses) required to calculate the U value statistic in a manner that allowed for a degree of certainty that the observed distribution was approximately representative of the "true" distribution. It was clear that having a large number of trials would result in the program taking a long time to complete, with boredom and distraction becoming a problematic possibility. However, it has been noted that U-value calculations require large amounts of data (Miller & Neuringer, 2000).

As the program aimed to incorporate lag schedules up to a lag 8 value, it was decided that four presses across two keys was not appropriate as this only allowed 16 possible responses. It would be difficult to fulfil a lag 8 schedule (where each

sequence must differ from the previous 8 sequences) under these conditions. Therefore, it was decided that four keys would be used. When using sequences of four presses across four keys, there are 256 possible sequences ($4^4$). Although there are no set criteria for determining the response number for the U-value statistic, a "minimum of four responses per alternative" has been suggested (G. Jensen, personal communication, 4th April, 2010). This would result in 1024 trials being required for each lag value in order to calculate a U-value for each. This was deemed impractical for the current purpose.

An alternative was decided upon that limited the number of possible sequences. It involved four presses and four keys, but each key could only be used once per sequence, so that the variability came from the order in which they were pressed. This reduced the number of possible sequences to 24. As the last response in such a sequence is a forced choice (as there is only one key remaining), this requirement was removed to decrease response effort, resulting in a "4 choose 3" permutation task. In this way, allowing for four trials per possible sequence resulted in 96 trials per lag value. See Table 4.3 for a list of all possible sequences of 3 presses across the four keys, A, B, C and D.

| Sequence | | | |
|---|---|---|---|
| ABC | BCD | CDA | DAB |
| ACB | BDC | CAD | DBA |
| ABD | BAC | CBA | DBA |
| ADB | BCA | CAB | DCB |
| ACD | BAD | CBD | DAC |
| ADC | BDA | CDB | DCA |

*Table 4.3:* List of all possible sequences of three presses across four keys.

The program consisted of 6 different conditions, where each condition consisted of 96 trials, resulting in a total trial number of 576 per participant who completed the program. Five of the conditions corresponded to a different lag value (1, 2, 4, 6, or 8). The sixth acted as a control condition. Within the program, participants were randomly assigned to one of four settings: A, B, C or D. Each setting presented the conditions in a different order, the purpose of which was to control for the order of presentation of the lag values. The aim of the study was not to investigate the effects of increasing the lag value incrementally across conditions, but rather to look at performances under different conditions reinforcing variability. Therefore, it was important to alter the order of presentation of conditions across participants. Table 4.4 describes the order in which the conditions were presented in each of the four settings.

| Setting | Condition 1 | Condition 2 | Condition 3 | Condition 4 | Condition 5 | Condition 6 |
|---|---|---|---|---|---|---|
| A | lag 1 | lag 2 | control | lag 4 | lag 6 | lag 8 |
| B | lag 2 | lag 4 | lag 8 | lag 1 | control | lag 6 |
| C | lag 1 | control | lag 4 | lag 8 | lag 6 | lag 2 |
| D | control | lag 2 | lag 1 | lag 8 | lag 4 | lag 6 |

*Table 4.4:* List of the four program settings with the order of conditions for each.

At the end of the program, a U-value was automatically calculated and stored in an output file on the laptop along with a list of all the sequences emitted by the participant across the six conditions.

***Description of the computer program.*** The participants were presented with an animated scene and an associated character in each level. A level was defined as the process of moving the animal, from the start point to the end point, where they reached a target item (for example, a monkey swinging from tree to tree to reach a bunch of bananas). Reinforcement consisted of the character moving progressively through the scene. The levels progressed from one to the next as the participants obtained reinforcement and several levels could be observed by the participants within the same lag condition.

The program consisted of 10 levels that were presented to participants in order as they progressed through the program. No level was associated with a particular lag value. Participants required different numbers of trials to reach the end of a level, depending on whether they fulfilled the lag criterion and obtained reinforcement or not. As each lag condition required a fixed 96 trials it was not practical to attach a lag value to a level.

In each of the six conditions 96 trials were required in order to calculate the U-value for that condition, regardless of whether the participant was progressing

through the levels (i.e., obtaining reinforcement). If the 96 trials were reached in the middle of a level, every second trial after that was reinforced until the end of the level was reached but these trials were not recorded by the program. This allowed for the fact that each time a new condition began it was also at a new level.

The participants were not aware of the schedules in place at any time so they were responding solely to the consequences of the schedule (i.e., whether the character moved or not following a particular sequence of presses). The object in each level was to move the character across 10 obstacles to reach a location or item. In order to make the character move, participants were required to emit sequences of three presses across four available keys where each key could only be pressed once. Each time a key was pressed a tone sounded. When a participant pressed a key more than once in a sequence, a red X appeared on the screen, there was no tone and the last press was not recorded by the computer. The participant was required to press another key and if it had not been pressed already then it completed the sequence. After every trial (three presses) a box containing the word “wait” in red letters appeared on the screen for 1 second. While the box was on the screen, any keys pressed did not result in a tone and they were not recorded by the computer. This ensured that the participants were responding in sequences of three presses and not just pressing keys continuously.

As well as the character moving forwards, the computer offered written feedback (e.g., “well done”; “great”) for every reinforced trial. At the end of each level, fanfare music played and “you did it” appeared across the screen in large letters. Figures 4.2, 4.3 and 4.4 represent the scenes and characters for three of the levels.

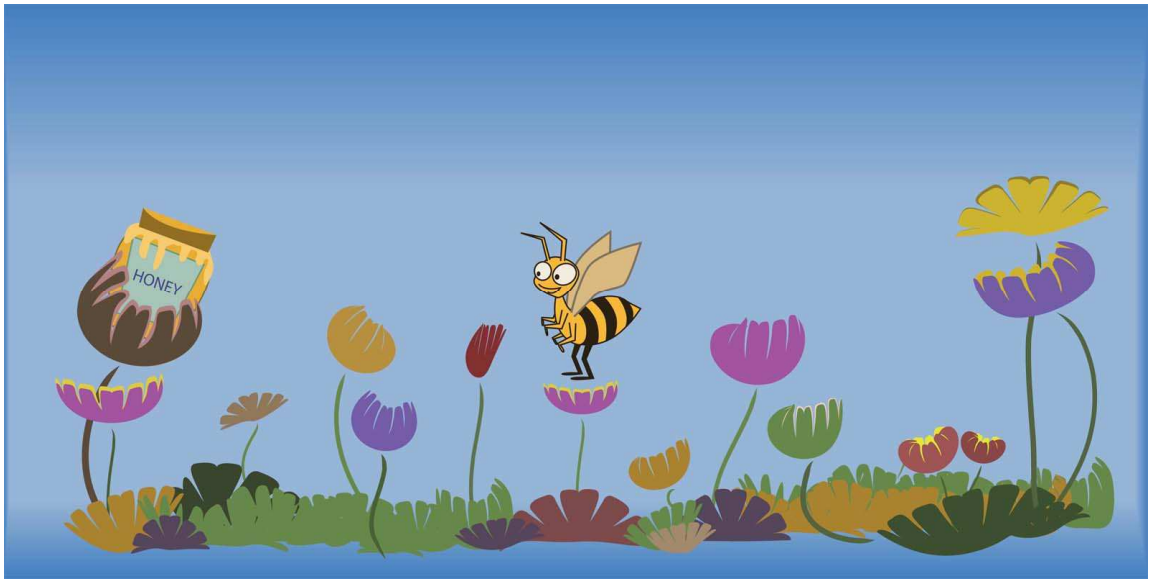


*Figure 4.2:* One level of the program: bee travelling along the flowers to reach honey.

*Figure 4.3:* One level of the program: mouse looking for cheese.

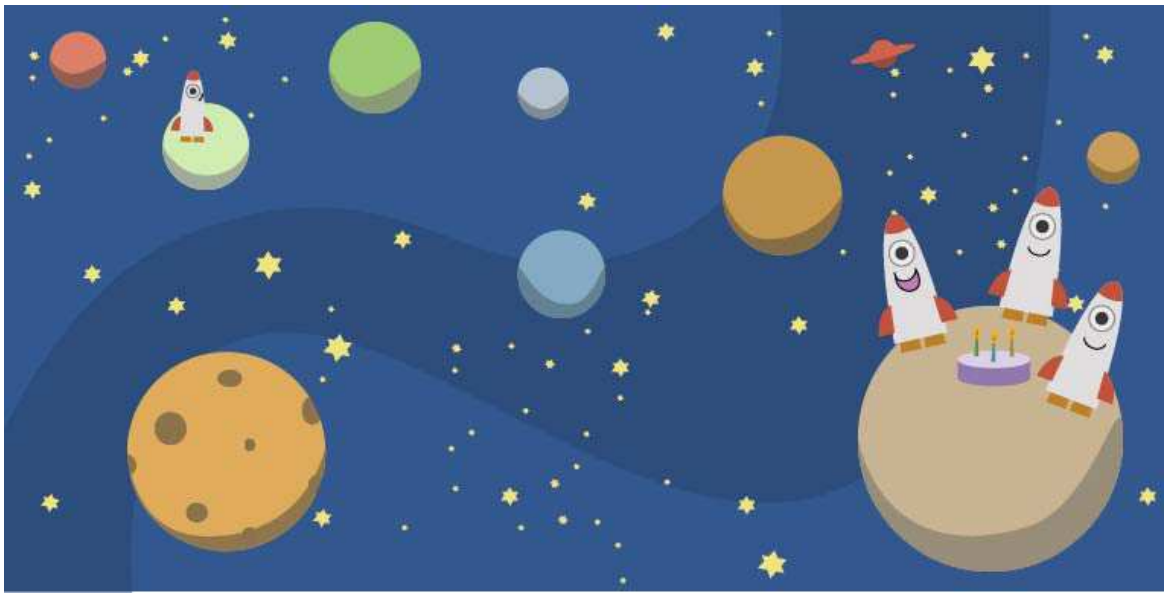

*Figure 4.4:* One level of the program: rocket travelling across planets to reach a birthday party.

**Measures**

**Autism Diagnostic Observation Schedule- Generic (ADOS-G, Lord et al., 2000)**

The ADOS-G is a semi-structured, standardised assessment of social, communication and play skills and the imaginative usage of materials used for the diagnosis of autism spectrum disorders. Social and communication opportunities are designed in a way that will elicit social and communicative behaviours. Play situations are set up in order to observe a range of role plays and imaginative activities. The ADOS-G consists of structured activities and materials and less-structured interactions that allow for a comprehensive evaluation of the individual's performance in the areas of social skills, communication, imaginative use of objects and other behaviours relevant to an ASD diagnosis (Lord et al., 2000). The ADOS-G was developed from 2 similar diagnostic tools, the Autism Diagnostic Observation

Schedule (Lord et al., 1989) and the Pre-Linguistic Autism Diagnostic Observation Scale (DiLavore, Lord & Rutter, 1995).

The ADOS-G consists of 4 modules that take approximately 30 minutes to administer. Only one module is administered at a given point in time and each module is suitable for children or adults at different levels of language and development. In the previous versions of the instrument, sufficient controls for the effects of language were not provided. Individuals with intellectual disability with or without ASD have been found to show greater performance when language demands are lower, relative to their ability. This meant that previous versions were at risk of over-diagnosing ASD in individuals with poor language ability; and under-diagnosing ASD in individuals whose language ability was more advanced than those for whom the scales were devised. The ADOS-G utilises different modules in order to control for the biasing effects of language by providing different tasks and coding in the different modules - keeping the diagnosis separate from the language abilities of the individual.

The ADOS-G provides a classification of either autism or autism spectrum disorder (ASD). For a diagnosis of autism, the individual must exceed the specified thresholds on the social domain, the communication domain and social-communication total. If they are outside these thresholds, they may meet the thresholds for ASD. These are lower than those for autism as it is considered a less severe classification.

The ADOS-G was found to have excellent reliability (Lord at al., 2000). Intraclass correlations were conducted between domains on the 4 modules. Correlations ranged from .88 to .97 for the social domain, .74 to .90 for the communication domain and from .84 to .98 for the social-communication total.

Interrater agreement for diagnostic classifications ranged from 84% to 100%. Test-retest reliability was also assessed and the stability of the social and communication domains and the social-communication total was found to be excellent (.78, .73 and .82 respectively).

Item-total correlations were conducted and they ranged from .62 to .88 for the communication domain and from .52 to .90 for the social domain. Chronbach's alphas were conducted to assess for internal consistency. These were found to be very high for the social domain (.85 to .92 for each module) and slightly lower for the communication domain (.74 to .84).

The test has been shown to demonstrate good reliability and validity as a diagnostic instrument. As autism is a complex disorder to diagnose, other clinical indications will always be considered such as the presence of restricted repetitive behaviours. However, for research purposes, it offers a reliable indication of the severity of ASD symptoms and a useful profile of the participants.

**Setting**

The participants in Group 1 either completed the program in their school or in their home. For seven of the participants the study was conducted in their schools in a quiet room adjacent to their classroom which they were familiar with. Only the experimenter and the child were in the room together with an open door. The participant sat at a table with the laptop and Ergodex® while the experimenter sat approximately one metre away. Three of the participants completed the program in their own homes. In these cases, the study took place in a quiet room with no distractions such as TV or toys. As before, only the experimenter and the participant were in the room with an open door. The experimenter was seated one metre away

from the participant who sat at a table with nothing on it except for the study apparatus.

The participants in Group 2 completed the program in a quiet room in their school. The set-up was identical to that of Group 1.

**Procedure**

The ADOS-G (Lord et al., 2000) was administered to each participant in Group 1 in either their school (n = 7) or their home (n = 3) under the conditions described above. The computer program was administered within two weeks of the ADOS-G for all participants in this group. The ADOS-G and the computer program were administered in the same setting for each child.

Prior to beginning the program, the participant was brought to the room and given the following verbal instructions:

*"Here is a new computer game for you to play. You can press these buttons to make the animals move across the screen. They are all looking for something and you can help them to find it. Press any three buttons in a row and then you must wait a second to see what happens. I'll help you for the first level and then you can do it by yourself."*

The experimenter logged the participant onto the program using a unique ID number. A practice level, utilising a lag 0 schedule of reinforcement was implemented first. This will be referred to as the "Training Phase". Here, every trial obtained reinforcement in order to ensure that the participants were interested in the program. It also served to train the participants to respond in sequences of three presses with an inter-trial interval. During the practice trial, if the participant did not show the required inter-trial interval, the experimenter pointed to the "wait" box on

screen and said “you must wait a second before the next three presses”. After each trial the experimenter pointed to the written praise on the screen and read it out so that the participant would attend to it during the entire program. If the participant pressed the same key twice the experimenter pointed to the red X on screen and said “you can only press each button once in each set of three presses”. It was decided to prompt this during the Training Phase so that the participants learned as they interacted with the program.

Following completion of the Training Phase the Testing Phase began. Here, participants were randomly assigned to one of four settings (A, B, C or D). These settings corresponded to different sequences of lag schedules (see Table 4.4). In the Testing Phase, the participant interacted with the game in the same way as the Training Phase. Typically, participants cycled through the 10 levels two-three times before they completed the entire 576 trials across the six conditions of the Testing Phase. Each time they reached the end of the 10$^{th}$ level the experimenter logged them out of the program, thanked them for their participation and told them they could play again tomorrow. It was decided that the entire program was too long to complete in one day as it took approximately two hours in total. In order to avoid a satiation effect, the program was completed across two or three consecutive days.

If the participant requested a break in the course of the study, the experimenter could pause the game and then log them on again after the break using their ID number. When the program was finished (i.e., 576 trials had been recorded across the six conditions) a screen appeared with the following text:

*“Well done, you have finished the game!”*

**Data Analysis**

U-values were compared across conditions in order to identify if there was a significant effect of the lag value on variability in responding. U-values were also compared across the two groups to determine if there was a significant effect for group on variability in responding on this task.

## 4.3. Results and Discussion

Table 4.5 contains the U-values obtained in each condition by participants in Group 1 along with the mean U-value for each condition for the 10 participants in this group.

| **Group 1** | | **U-value** | | | | | |
|---|---|---|---|---|---|---|---|
| **Participant Number/ Setting** | | **Control** | **lag 1** | **lag 2** | **lag 4** | **lag 6** | **lag 8** |
| 1 | C | 0.592813 | 0.642345 | 0.748634 | 0.598438 | 0.594293 | 0.64144 |
| 2 | D | 0.279843 | 0.345213 | 0.287722 | 0.896598 | 0.943686 | 0.573177 |
| 3 | C | 0.308751 | 0.520652 | 0.445302 | 0.53503 | 0.485445 | 0.603042 |
| 4 | D | 0.191487 | 0.0000 | 0.437521 | 0.031864 | 0.206419 | 0.204524 |
| 5 | A | 0.872429 | 0.895221 | 0.898224 | 0.892566 | 0.878317 | 0.928573 |
| 6 | A | 0.821843 | 0.938869 | 0.864586 | 0.812814 | 0.826624 | 0.884428 |
| 7 | B | 0.781524 | 0.764475 | 0.458593 | 0.739301 | 0.814078 | 0.702826 |
| 8 | A | 0.559266 | 0.434939 | 0.656062 | 0.704402 | 0.839627 | 0.860318 |
| 9 | C | 0.466084 | 0.549926 | 0.891932 | 0.923217 | 0.91453 | 0.950743 |
| 10 | B | 0.492069 | 0.828779 | 0.860142 | 0.866631 | 0.843982 | 0.797888 |
| **Mean U-value** | | 0.536611 | 0.592042 | 0.654872 | 0.700186 | 0.7327 | 0.714696 |

*Table 4.5:* U-values for the participants in Group 1 across all conditions, together with mean U-value for each condition.

Table 4.6 contains the U-values obtained in each condition by participants in Group 2 along with the mean U-value for each condition for the 10 participants in this group.

| **Group 2** | | **U-value** | | | | | |
|---|---|---|---|---|---|---|---|
| **Participant Number/ Setting** | | **Control** | **lag 1** | **lag 2** | **lag 4** | **lag 6** | **lag 8** |
| 11 | C | 0.848607 | 0.847276 | 0.818412 | 0.853492 | 0.894714 | 0.875107 |
| 12 | B | 0.902725 | 0.880104 | 0.885469 | 0.927196 | 0.841349 | 0.937243 |
| 13 | C | 0.89732 | 0.795937 | 0.763755 | 0.872883 | 0.910925 | 0.864833 |
| 14 | D | 0.882687 | 0.88046 | 0.905255 | 0.815912 | 0.815039 | 0.913938 |
| 15 | D | 0.932173 | 0.827014 | 0.924413 | 0.900941 | 0.918056 | 0.874965 |
| 16 | A | 0.646649 | 0.777211 | 0.821215 | 0.803835 | 0.828575 | 0.704622 |
| 17 | A | 0.811881 | 0.847598 | 0.9017 | 0.806881 | 0.90692 | 0.870169 |
| 18 | A | 0.548768 | 0.670869 | 0.715672 | 0.882844 | 0.921311 | 0.923339 |
| 19 | D | 0.83761 | 0.743078 | 0.700565 | 0.870586 | 0.931969 | 0.812242 |
| 20 | B | 0.887605 | 0.911983 | 0.907818 | 0.905051 | 0.913675 | 0.89024 |
| **Mean U-value** | | 0.81903 | 0.819153 | 0.834427 | 0.863962 | 0.888253 | 0.86667 |

*Table 4.6:* U-values for the participants in Group 2 across all conditions, together with the mean U-value for each condition.

**Descriptive Statistics**

Table 4.7 shows the mean U-values and standard deviations for Group 1, Group 2 and Group 1/Group 2 combined across the six conditions. From the table, it is clear that Group 1 showed lower variability during the control condition when reinforcement was provided independent of variability. Interestingly, both groups demonstrated the highest variability when the lag 6 schedule was in effect. However, Group 2 had consistently high variability across the six conditions, showing little differentiation in the presence of the changing lag schedules.

| Condition | Group | Mean Uvalue | Standard Deviation | N |
|---|---|---|---|---|
| **Control** | 1 | .5357 | .2349 | 10 |
| | 2 | .8196 | .1241 | 10 |
| | Total | .6777 | .2338 | 20 |
| **lag 1** | 1 | .5921 | .2874 | 10 |
| | 2 | .8181 | .0729 | 10 |
| | Total | .7051 | .2347 | 20 |
| **lag 2** | 1 | .6549 | .2297 | 10 |
| | 2 | .8344 | .0837 | 10 |
| | Total | .7447 | .1918 | 20 |
| **lag 4** | 1 | .7001 | .2691 | 10 |
| | 2 | .8640 | .0432 | 10 |
| | Total | .7820 | .2055 | 20 |
| **lag 6** | 1 | .7327 | .2336 | 10 |
| | 2 | .8883 | .0429 | 10 |
| | Total | .8105 | .1819 | 20 |
| **lag 8** | 1 | .7147 | .2253 | 10 |
| | 2 | .8667 | .0670 | 10 |
| | Total | .7907 | .1796 | 20 |

*Table 4.7:* Descriptive statistics: Mean U-values and standard deviations for Group 1, Group 2 and Group 1/Group 2 combined for each of the six conditions.

Figures 4.5 and 4.6 illustrate the mean U-values for each condition for Groups 1 and 2 respectively. While the trends are identical across the two groups, the effect is much clearer for Group 1, indicating that the schedule in place had a greater effect on their variability. One possible explanation for this is the high level of variability demonstrated by Group 2 during the control condition, i.e., they responded variably on this task without the lag requirement. However, in the presence of the lag schedule, variability was increased. For both groups, the highest U-values were observed in the higher lag conditions (lag 4, lag 6 and lag 8).

It is interesting to note that the lag 6 schedule produced higher variability than the lag 8 schedule. Of interest to the experimenter when designing the study

was whether the higher lag requirements would actually result in lower variability. It is thought that when the lag value becomes too high an extinction effect on variability occurs as the magnitude of reinforcement drops and the response effort increases. This is a likely explanation for the outcome of the current study and it points to the consideration of a possible ceiling effect when using such schedules with human participants.

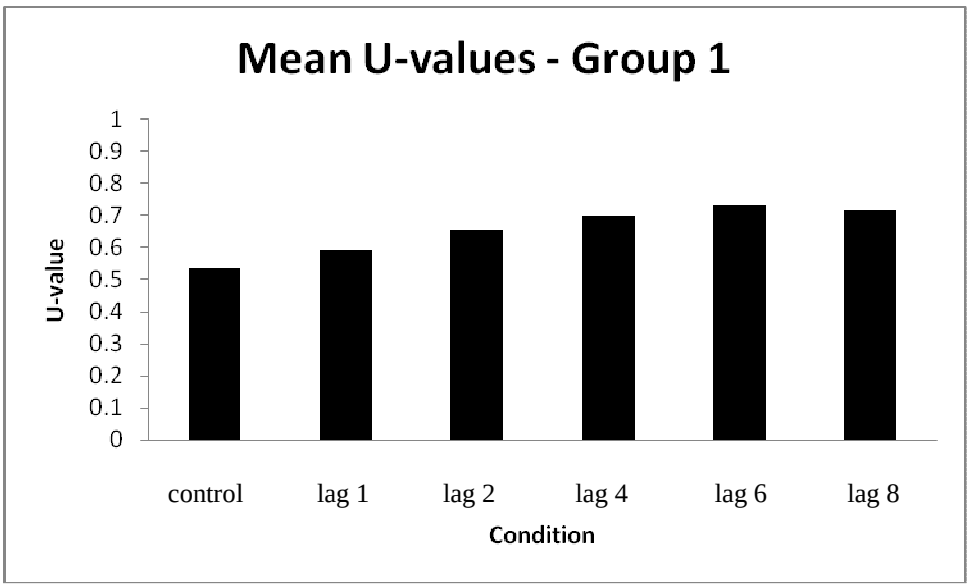


*Figure 4.5:* Mean U-values for Group 1 across the six conditions.

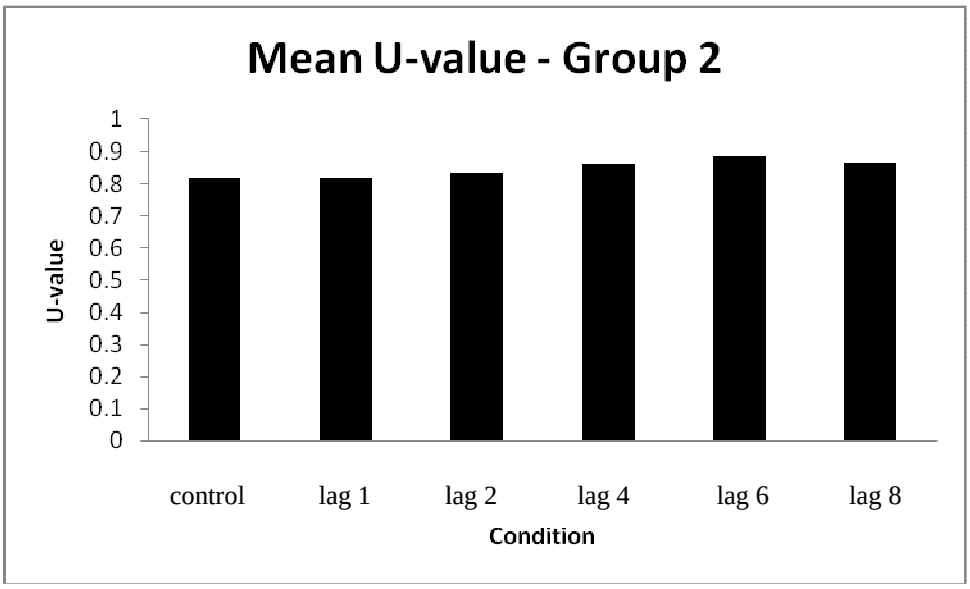


*Figure 4.6:* Mean U-values for Group 2 across the six conditions.

**Analysis of Group Data**

**Within Subject Effects.** A mixed ANOVA was conducted to investigate the within subject effects of the lag schedule on U-value. Mauchley's Test of Sphericity was significant, indicating that the assumption of sphericity had been violated. The repeated measures ANOVA with a Greenhouse-Geisser correction determined that there was a significant main effect for schedule on U-value ($F(2.67, 48.05) = 4.407$, $p = 0.01$, partial $\eta^2 = 0.20$). Cohen (1988) discussed the magnitude of effect sizes when using partial $\eta^2$, describing 0.09 as a medium effect size and 0.25 as a large effect size. The effect size of 0.2 obtained in the current study closely approaches a large effect size.

The schedule *x* group interaction effect was not significant ($F(2.27) = 1.092$, $p = 0.357$, partial $\eta^2 = 0.06$) indicating that there was no interaction effect of group and schedule in this sample.

**Between-Subject Effects.** There was a significant effect for group on U-value in this study ($F(1,18) = 8.00$, $p = 0.011$), with the participants in Group 2 demonstrating significantly higher variability than those in Group 1.

**Post hoc Tests.** To further investigate the main within subject effect, post hoc tests were conducted. Five sets of paired sampled t-tests were used to identify if there were significant differences in U-values between the control condition and the five lag reinforcement conditions utilised in this study. The independent variable was the lag value of the schedule and the dependent variable was the U-value. The results indicated that there was no significant difference in U-values between the control condition and the lag 1 condition, $t(19) = -.99$, $p = .333$ or between the control condition and the lag 2 condition, $t(19) = -1.757$, $p = .095$ or between the control condition and the lag 4 condition, $t(19) = -2.340$, $p = .03$. However,

significant differences were observed between the control condition and the lag 6 condition, $t(19) = -3.069$, $p = .006$ and between the control condition and the lag 8 condition, $t(19) = -3.091$, $p = .006$. The results indicated that the higher lag values were more effective in increasing response variability relative to a control condition.

It is clear that the difference between the control condition and the lag 4 condition was significant at the 0.05 level but when a Bonferroni correction was applied it did not reach the required level. This difference may be considered borderline but there was a robust effect between the control condition and the lag 6 condition and the control condition and the lag 8 condition.

**Analysis of Individual Participant Data**

Following statistical analysis it was decided to conduct an additional single case analysis. The rationale for this was that no research study had examined the effects of a comprehensive series of lag schedules on variability in responding before now. Furthermore, it was thought that there would be merit in investigating the precise effects of the different lag schedules on the neurotypical participants and those with ASD. The use of visual displays of individual data sets was also desirable as patterns of responses would be visible and any effects of the order of presentation of the lag schedules would emerge clearly upon inspection of such displays. The U-value statistic was especially useful here as it allowed for comparison across conditions, program settings (order), participants and groups. Hence, it was decided that in order to fully utilise the wealth of available data, a single case analysis was warranted.

**Group 1.** There was considerable variation in how the participants in Group 1 responded to the various conditions of the computer program. Figures 4.7 to 4.16

below represent the individual scores across the six conditions along with the order in which the conditions were presented for each participant. In general, participants achieved the highest U-values when the lag requirement was increased. However, there are some notable exceptions.

Figure 4.7 represents the U-values obtained by Participant 1 across the six conditions. The U-values obtained by this participant in the control, lag 4 and lag 6 conditions are very similar (0.5928, 0.5984 and 0.5943 respectively). Interestingly, this participant demonstrated higher variability in the lag 1 condition (U-value = 0.6423), which was presented first, and this was almost identical to the lag 8 condition (U-value = 0.6414). Participant 1 achieved the highest U-value in the lag 2 condition (0.7486). This participant did not show differentiation between conditions as would be expected due to the changing lag value.

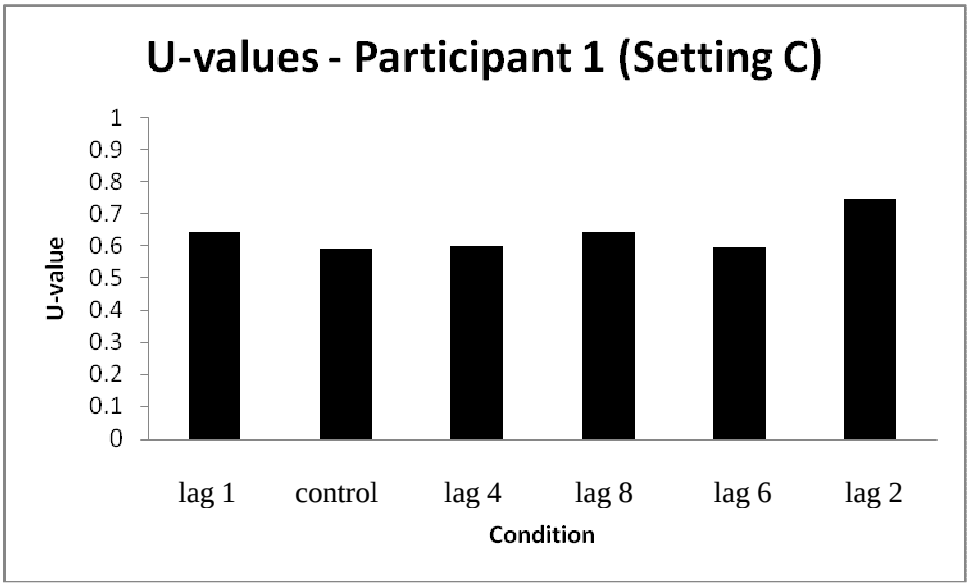


*Figure 4.7:* U-value scores for Participant 1 across the six conditions.

Participant 2 demonstrated U-values that were low and very similar across the control, lag 1 and lag 2 conditions (0.2798, 0.3452 and 0.2877 respectively). When the lag 8 schedule was implemented, the U-value increased to 0.5732. The value

further increased to 0.8966 in the lag 4 condition which represented a high level of variability in responding. This value further increased in the presence of the lag 6 schedule to 0.9437 (see Figure 4.8).

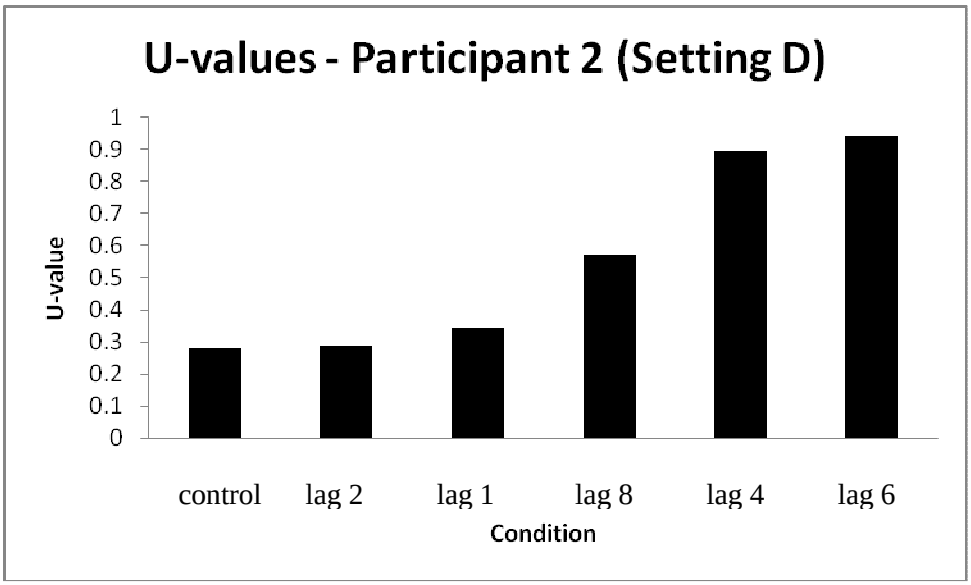


*Figure 4.8:* U-value scores for Participant 2 across the six conditions.

Participant 3 began the computer program on a lag 1 schedule of reinforcement. This participant demonstrated the lowest U-value in the control condition (0.3088) which was presented after the lag 1 condition. This indicates that the each of the five lag schedules increased variability in responding for this participan relative to the control condition. Of note is the high U-value evident in the lag 8 condition (0.6030). For this participant, the lag 8 schedule was readily fulfilled and yielded the highest variability. However, Participant 3 did not reach the high U-values that many of the other participants did, irrespective of what schedule was in place (see Figure 4.9).

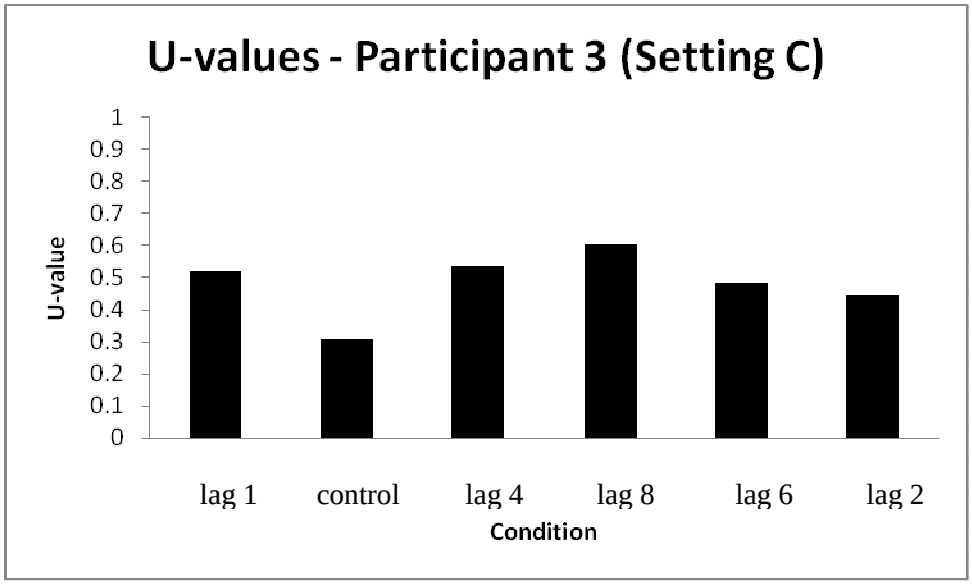


*Figure 4.9:* U-value scores for Participant 3 across the six conditions.

Participant 4 obtained a U-value of 0 in the lag 1 condition as he only emitted one sequence across the 96 trials showing no variation across trials. This was the only time a U-value of 0 was obtained in the course of the study. The highest U-value for this participant was in the lag 2 condition, where he obtained a U-value of 0.4375. In the presence of the higher lag requirement, Participant 4 did not demonstrate increased variability but rather returned to an almost identical U-value to that of the control condition. U-values for Participant 4 are presented in Figure 4.10.

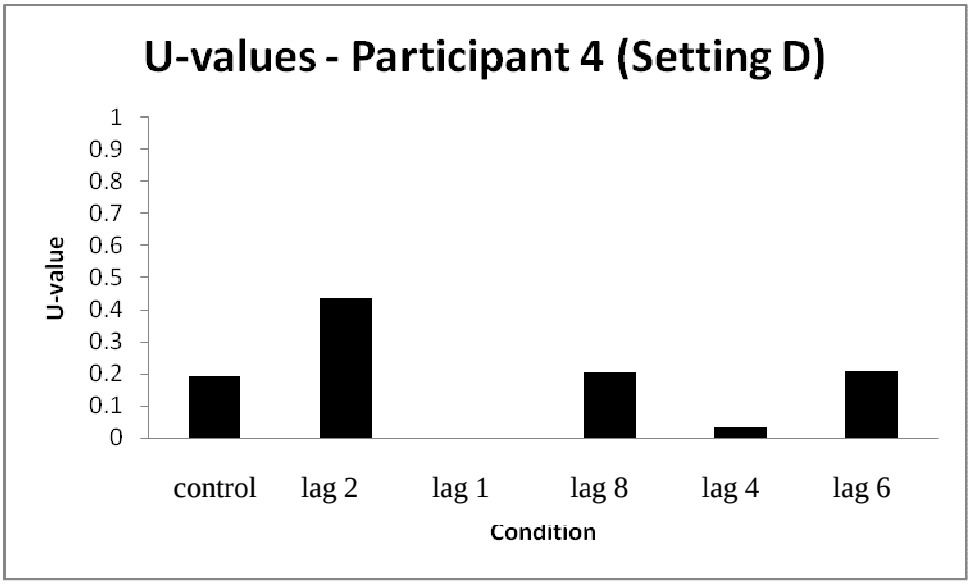


*Figure 4.10:* U-value scores for Participant 4 across the six conditions.

Figure 4.11 represents the U-values obtained by Participant 5. This participant achieved high U-values across the control, lag 1, lag 2, lag 4 and lag 6 conditions and they remained within an extremely narrow range (0.8725 - 0.8982). Contrary to the majority of other participants, this participant achieved the highest U-value in the lag 8 condition (0.9286) indicating that the increased lag requirement did not result in an extinction effect.

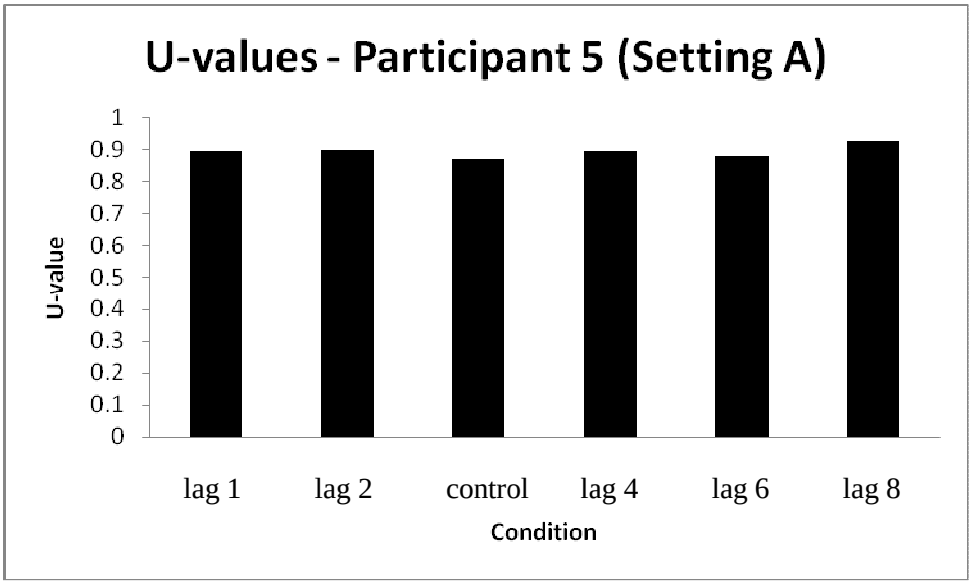


*Figure 4.11:* U-value scores for Participant 5 across the six conditions.

The variability in responding for Participant 6 did not appear to be under the control of the lag schedules of reinforcement. This participant achieved U-values that were almost identical across the six conditions (0.8128 – 0.9389), indicating no discernable difference between the control condition and the lag reinforcement conditions (see Figure 4.12). Participant 6 indicated dissatisfaction with the computer program several times in the course of the study and this may have contributed to his lack of responsivity to the different conditions.

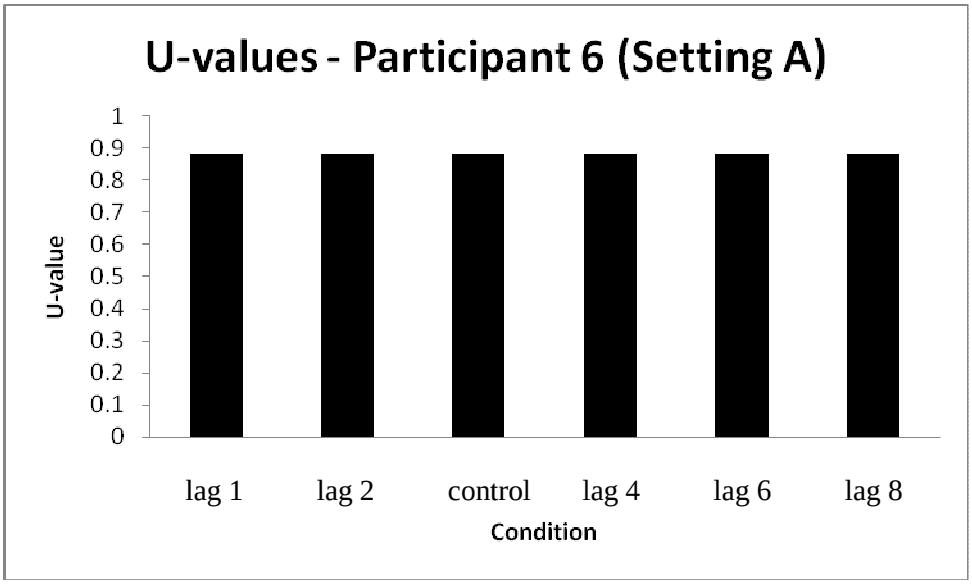


*Figure 4.12:* U-value scores for Participant 6 across the six conditions.

The U-values achieved by Participant 7 are represented in Figure 4.13. Participant 7 appeared to show a learning effect of the lag schedule as the lowest U-value was achieved in the lag 2 condition (0.4586). This increased to 0.7645 in the lag 4 condition and this increase was essentially maintained across the next 4 conditions including the control condition (0.7815). The highest U-value was recorded in the lag 6 condition which was presented last.

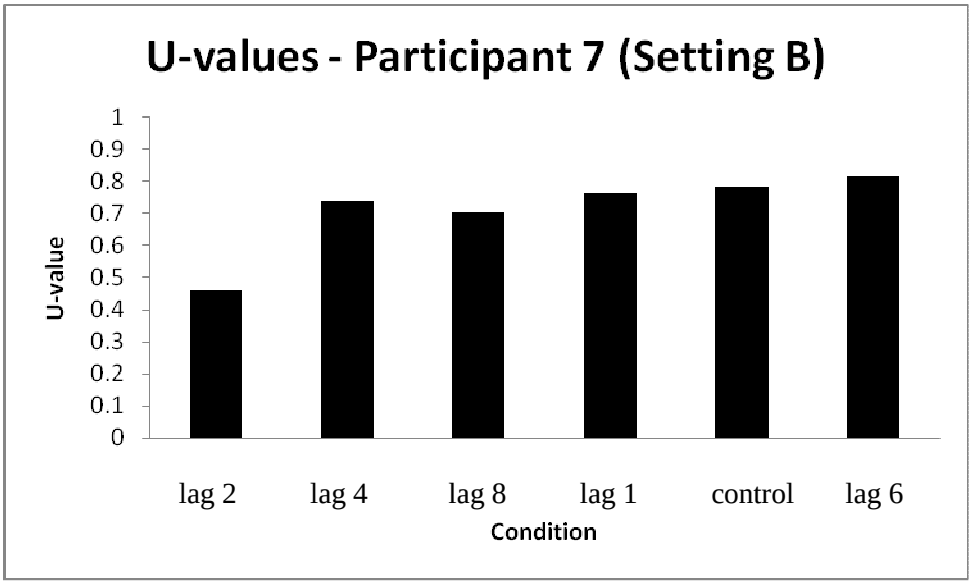


*Figure 4.13:* U-value scores for Participant 7 across the six conditions.

Figure 4.14 represents the U-values achieved by Participant 8. Interestingly, the lag 2 condition yielded much higher variability than the lag 1 condition (U-value = 0.6560 vs 0.4349). The U-value for the control condition was 0.5593. The U-value increased across the final three conditions to 0.8603 in the lag 8 condition. This represents a marked increase in variability in responding relative to the control condition indicating that this participant was responding to the changing lag schedules throughout the computer program.

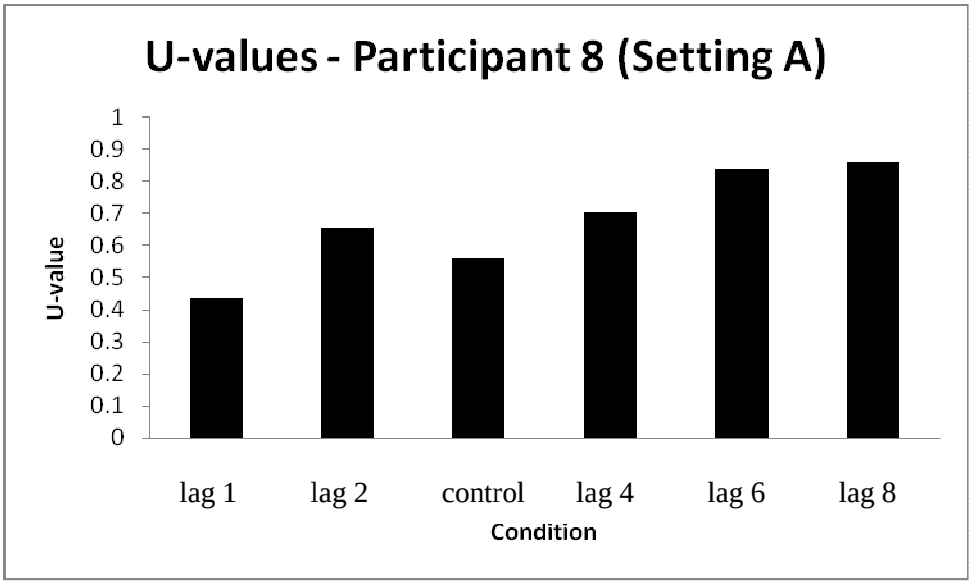


*Figure 4.14:* U-value scores for Participant 8 across the six conditions.

Participants 9 and 10 both demonstrated low to moderate variability in the control condition (with U-values of 0.4661 and 0.4921 respectively) and responded to the lag requirements with greatly increased U-values. Participant 9 demonstrated high variability in responding in the lag 2 (U-value = 0.8919), 4 (U-value = 0.9232) and 6 (U-value = 0.9143) conditions but the highest level of variability was observed in the lag 8 condition (U-value = 0.9507) (see Figure 4.15). Participant 10 demonstrated a marked increase in variability in the lag 1 (U-value = 0.8288), lag 2 (U-value = 0.8601), lag 4 (U-value = .8666), lag 6 (U-value = 0.8240) and lag 8 (U-value = 0.7979) conditions relative to the control condition with the highest U-value in the lag 4 condition, which was presented second, and the lowest in the lag 8 condition, which was presented third (see Figure 4.16). These data suggest the order did not have any effect on the variation shown by these participants and they were responding to the lag schedule in place in each condition. However, the fact that Participants 9 and 10 responded to the schedules differently indicates that some children with ASD may demonstrate higher variability in the presence of higher lag

values while others may respond more variably when the lag requirement is not increased past a lag 4.

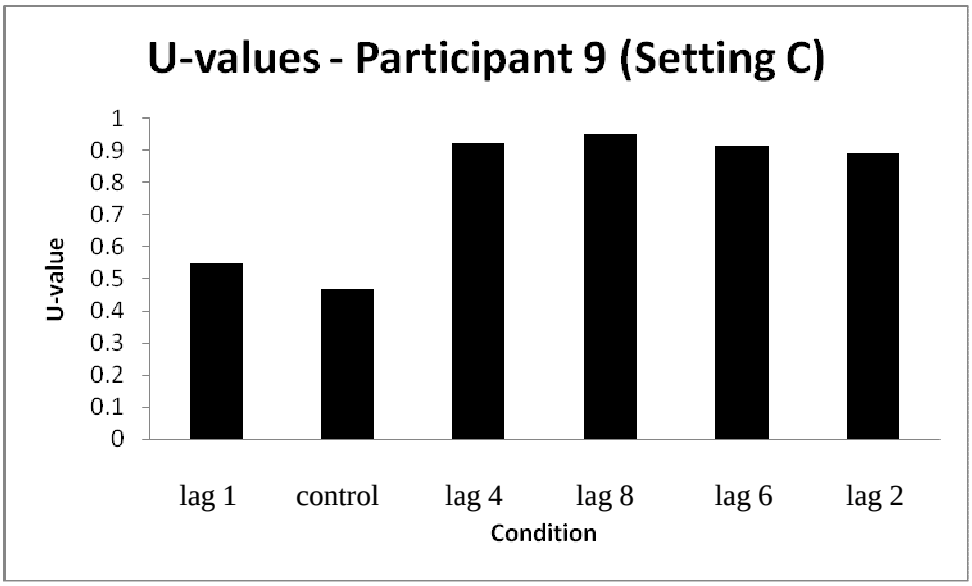


*Figure 4.15:* U-value scores for Participant 9 across the six conditions.

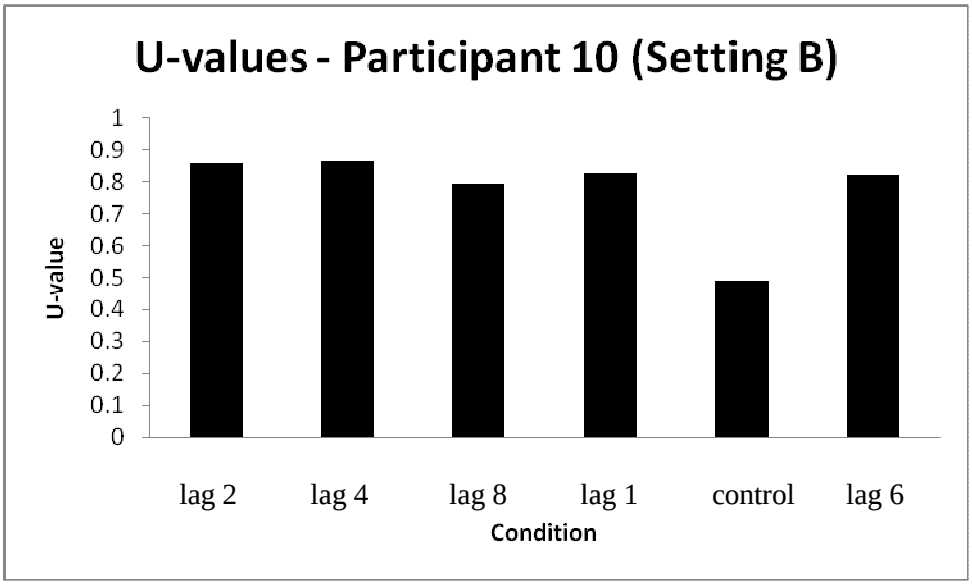


*Figure 4.16:* U-value scores for Participant 10 across the six conditions.

**Variability and Autism Severity.** The scores obtained by the participants in group 1 on the ADOS-R (Lord et al., 2000) were examined in order to identify if severity of autism as indicated by the ADOS-R was related to the level of variability

of the participants and their responsiveness to the lag schedules. Four of the participants (Participants 1, 2, 5 and 6) had low ADOS-R scores (total score of 8 or 9), indicating low autism severity and a disgnosis of autism spectrum rather than autism according to this measure. In the control condition, where reinforcement was independent of variability in responding, two of them achieved U-values greater than 0.8, which is a high level of variability. However, one of the four achieved a low U-value (0.2798) representing very low variability and the other one achieved a U-value of .5928, indicating a moderate level of variability. Three of the four achieved U-values greater than 0.9 in the course of the study while the other one maintained a moderate level of variability throughout.

Participants 3, 8 and 9 showed ADOS-R scores in the range of a diagnosis of autism but their scores can be considerate moderate (total ADOS-R scores of 12, 13 and 14 respectively). All three achieved low to moderate U-values in the control condition (0.3 - 0.56). Only one of these (Participant 9) reached a U-value of greater than 0.9 in the course of the study. Participant 8 showed a steady increase in U-value across the conditions and Participant 3 showed a modest increase in variability but stayed within a narrow range of U-values.

Participants 4, 7 and 8 had high ADOS-R scores (28, 19 and 20 respectively), indicating a higher degree of autism severity than the other participants. Participant 4, who had the highest ADOS-R score of all the participants in Group 1, also showed the lowest variability in the course of the study. This participant had the lowest U-value at the control condition (0.1915) but, more interestingly, he achieved a U-value of 0 in the lag 1 condition. This means that he emitted the same sequence 96 times and this was the only sequence he emitted in the whole condition. He achieved another extremely low U-value in the lag 4 condition (0.0319) and he only had one

score that could be considered moderate (U-value = 0.4375 in the lag 2 condition). It is unclear whether the severity of autism symptoms is related to the extremely low variability demonstrated by this participant across the entire study. Participant 7 scored a moderate-high U-value of 0.7815 in the control condition and he achieved a U-value of 0.8141 in the lag 6 condition. Participant 10 achieved a moderate U-value in the control condition (0.4921) but he demonstrated a marked increase in variability across all of the lag conditions (U-values ranging from 0.7979 to 0.8666).

The participant with the highest variability overall in Group 1 was Participant 9 who scored three U-values greater than 0.9. Participant 9 was in the moderate range of the ADOS-R scores. Thus, while there is some evidence that greater autism severity was related to lower variability, the trend is not robust enough in the current sample to draw such conclusions.

**Group 2.** In the current study, participants in Group 2 demonstrated significantly higher variability in responding than those in Group 1. As with Group 1, the highest variability was noted in the presence of the higher lag values. Participant 11 achieved the highest U-value in the lag 6 condition (U-value = 0.8947) similar to the other participants in this study. However, the U-values across the other conditions were extremely similar, showing little change in u-value in response to the changing lag value. Participant 11 demonstrated high variability in the control condition (U-value = 0.8486), which may have contributed to this effect (see Figure 4.17) .

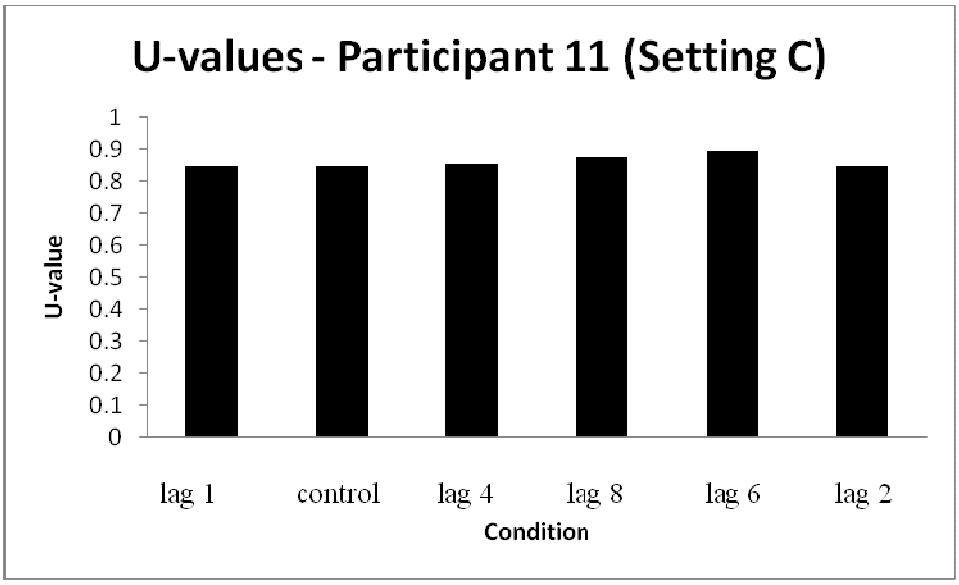


*Figure 4.17:* U-value scores for Participant 11 across the six conditions.

Participant 12 also demonstrated high variability in responding in the control condition. From Figure 4.18, we see that repsonding was slightly more variabile in the control condition (U-value = 0.9027) than in the lag 1 condition (U-value = 0.8801). Interestingly, the lowest variability was observed in the lag 6 condition (U-value = 0.8413) whilst the highest variability was observed in the lag 8 condition (U-value = 0.9372). The lag 6 condition was presented last for this participant and therefore a satiation effect may have occurred.

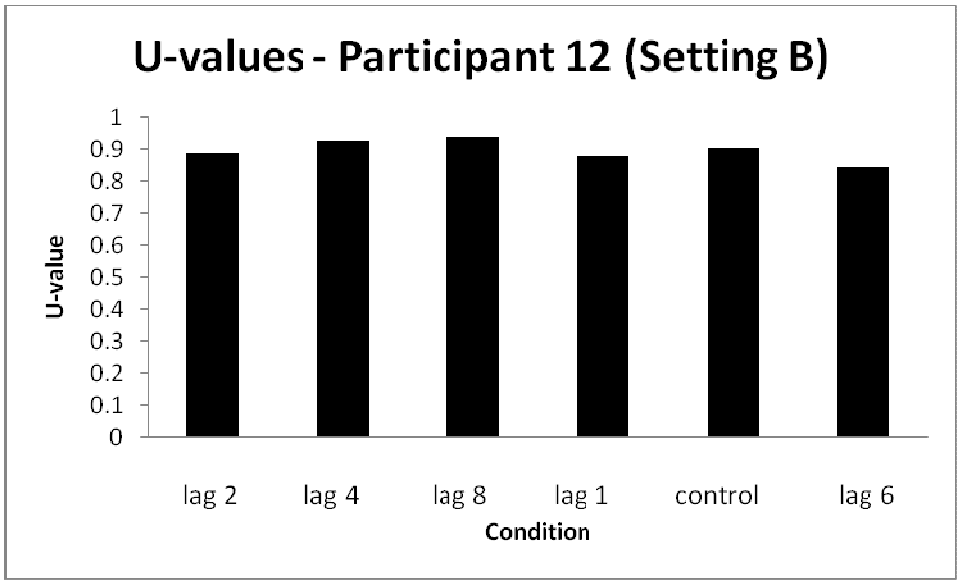


*Figure 4.18:* U-value scores for Participant 12 across the six conditions.

From Figure 4.19 it is clear that Participant 13 demonstrated a high level of variability in responding in the control condition (U-value = 0.8973). However, the U-value was lower than this in the lag 1 and lag 2 conditions (0.7959 and 0.7638 respectively). This outcome may indicate that the variability shown by Participant 13 was under the control of the schedule in place in each condition as the lag 1 and lag 2 criteria could be met without a high level of variability. In a lag 1 schedule the participant can receive reinforcement for every trial just by alternating between two responses while in a lag 2 schedule, cycling between three different responses will produce the same effect. Participant 13 showed increased variability in the lag 4 condition (U-value = 0.8729), while the highest U-value was observed in the lag 6 condition (U-value = 0.9109). Similar to other participants, in the lag 8 condition variability decreased with U-values converging toward the level observed in the control condition (U-value = 0.8648).

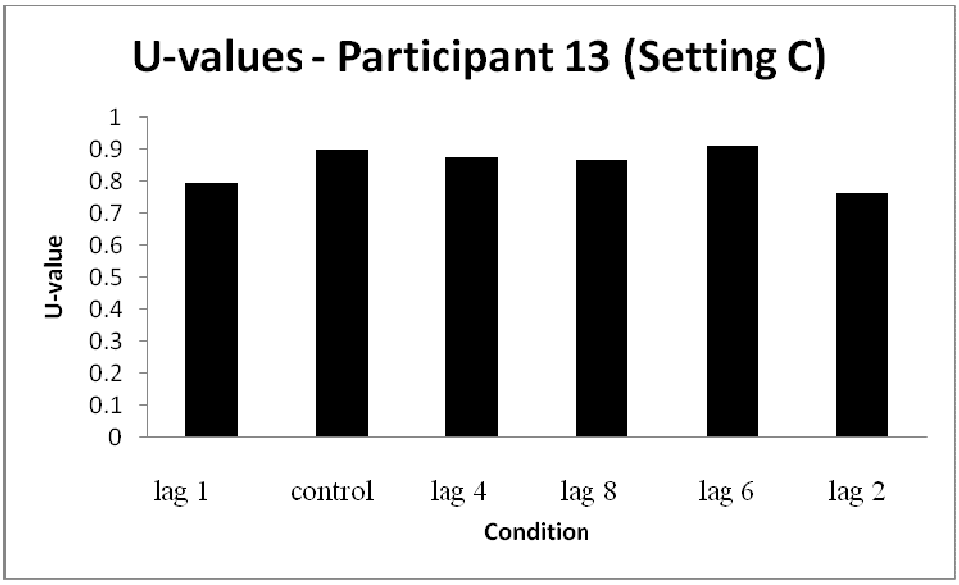


*Figure 4.19:* U-value scores for Participant 13 across the six conditions.

Participant 14 also demonstrated high variability in the control condition and did not show large differences in variability in responding as the lag requirement changed. However, the highest U-value was recorded in the lag 8 condition (0.9139), indicating that the schedule was having an effect on variability. Lower U-values were obtained in the lag 4 (0.8159) and lag 6 (0.8150) condition, which was contrary to the outcomes for most other participants (see Figure 4.20).

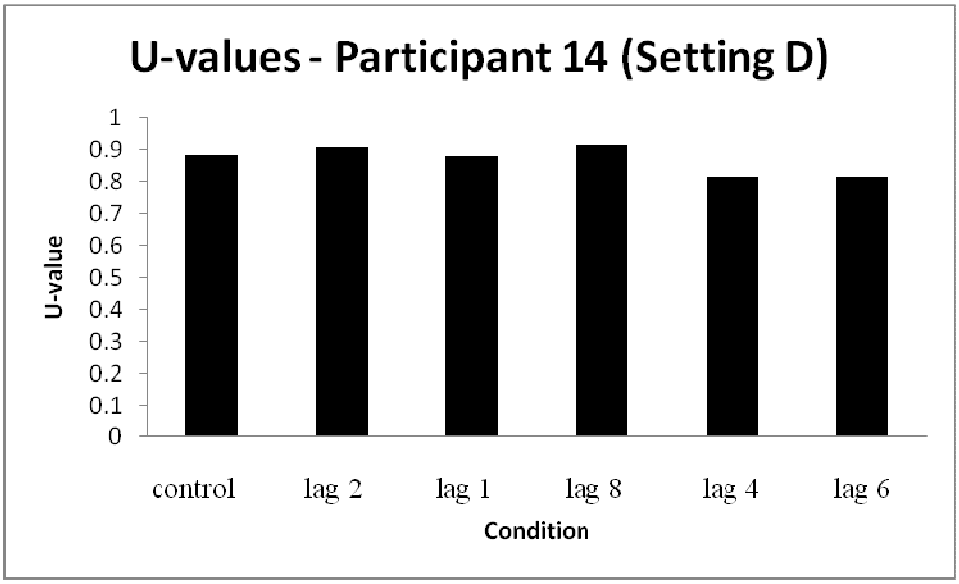


*Figure 4.20:* U-value scores for Participant 14 across the six conditions.

Figure 4.21 shows that Participant 15 demonstrated extremely high variability in responding in the control condition (U-value = 0.9322). This indicates that this participant had the ability to vary his responding and was inclined to do so, without reinforcement being contingent on it. What is interesting about this participnat's repsonse to the computer program is that he emitted lower variability in the lag 1 condition (U-value = 0.8370), indicating that this schedule can have an inhibiting effect on variability in those who behave variably in the absence of variation-dependent reinforcement.

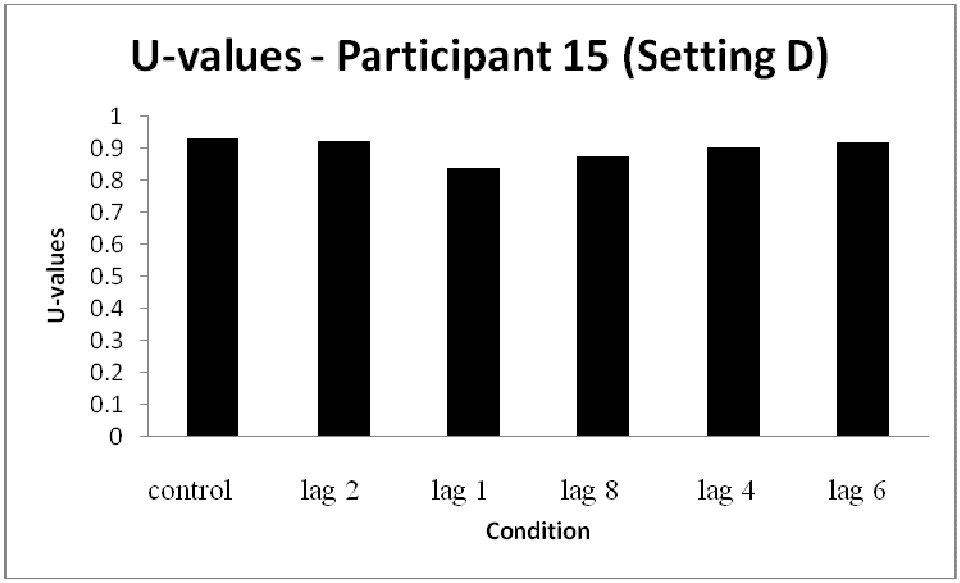


*Figure 4.21:* U-value scores for Participant 15 across the six conditions.

Participant 16 obtained the lowest U-value in the control condition (0.6466) so, although this was the third condition presented, when reinforcement was not contingent on variability, variation in responding decreased for this participant. U-values were similar across the lag 1 (0.7772), lag 2 (0.8212), lag 4 (0.8038) and lag 6 (0.8286). However, the U-value decreased to 0.7046 in the lag 8 condition,

indicating that this participant did not maintain the high level of variability when the lag value was increased past a lag 6 (see Figure 4.22).

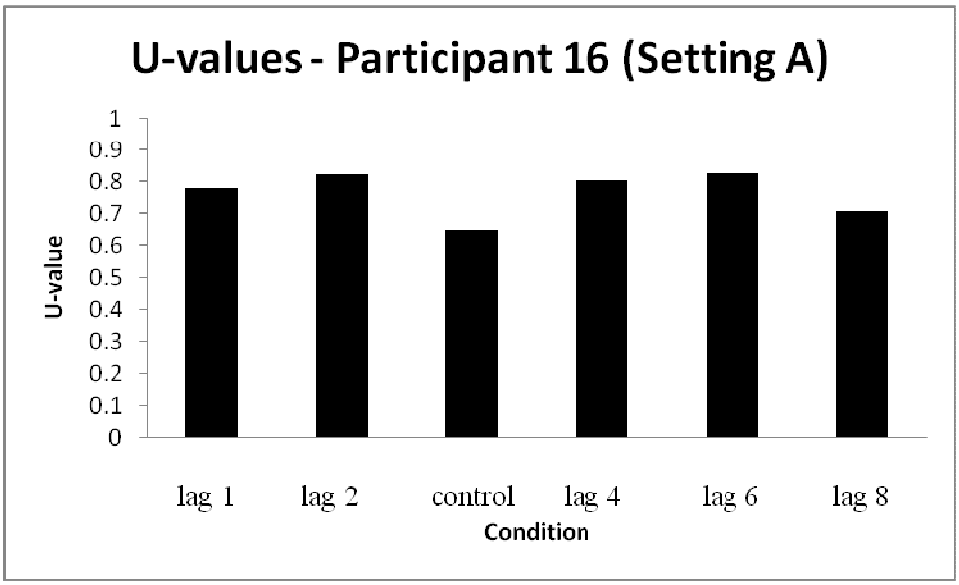


*Figure 4.22:* U-value scores for Participant 16 across the six conditions.

Figure 4.23 shows that Participant 17 responded in a similar way to Participant 16. Interestingly, both participants completed the conditions in the same order as both were on Setting A. Participant 17 had slightly higher U-values overall but he also showed a decrease in variability in the control condition althought this decrease was not as marked as it was for Participant 16 (U-value = 0.8119). The U-value was highest in the lag 6 condition (0.9069) as it was for the majority of participants and it decreased to 0.8701 in the lag 8 condition.

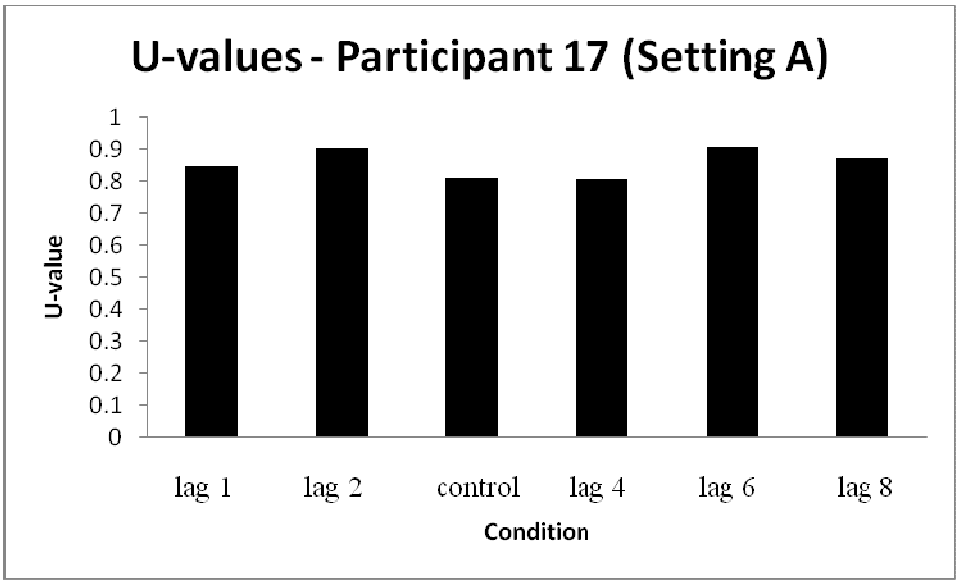


*Figure 4.23:* U-value scores for Participant 17 across the six conditions.

The data for Participant 18 are interesting as they clearly correspond with the lag schedule in place i.e., the higher the lag value the higher the U-value (see Figure 4.24). Participant 18 demonstrated a moderate level of variability in the control condition (U-value = 0.5488) and the U-value increased with the increasing lag requirement resulting in a high level of variability in the Lag 8 condition (U-value = 0.9233). Participant 18 interacted with the computer program on the A setting, where the lag value was increased systematically (i.e., from lag 1 to lag 8 in order). However, this setting did not influence the variability of Participants 16 and 17 as systematically as it did for Participant 18.

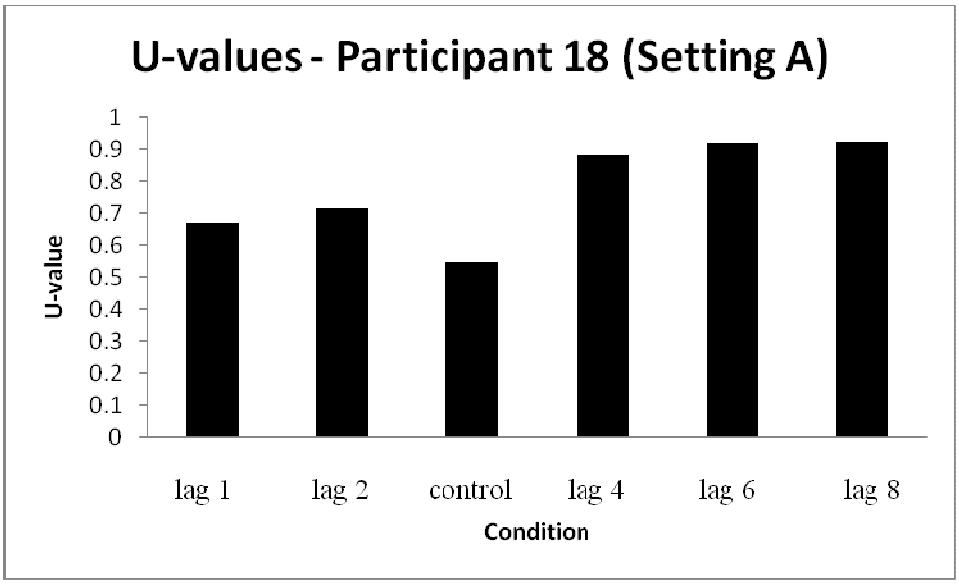


*Figure 4.24:* U-value scores for Participant 18 across the six conditions.

Figure 4.25 represents the U-values obtained by Participant 19. This participant demonstrated high variability in the control condition which was presented first (U-value = 0.8376). This decreased to 0.7006 in the next condition which was lag 2. The highest U-value was in the final condition which was lag 6 (0.9320). From Figure 4.25 we see that the U-value increased steadily across the five lag reinforcement conditions. Thus, it is unclear whether the specific lag requirement of each condition was influencing variability in responding or whether the participant was responding more variably as the program progressed as a function of a general requirement to “try something different” throughout the program.

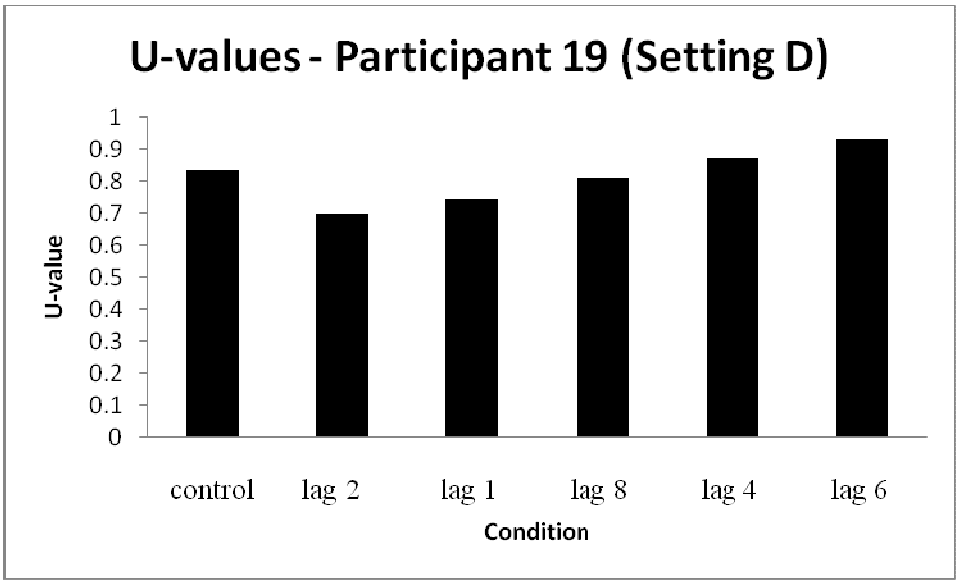


*Figure 4.25:* U-value scores for Participant 19 across the six conditions.

From Figure 4.26 it is clear that Participant 20 did not vary his behaviour in response to the lag schedule in place in the different conditions. One reason may be that this participant demonstrated very high variability in the control condition, independent of the lag schedule (U-value = 0.8876). Whilst the highest U-value was obtained in the lag 6 condition, similar to the general trend (0.9137), the range of U-values obtained by Participant 20 was so narrow (0.8876 – 0.9137) that the effect of condition on variability in responding cannot be assumed.

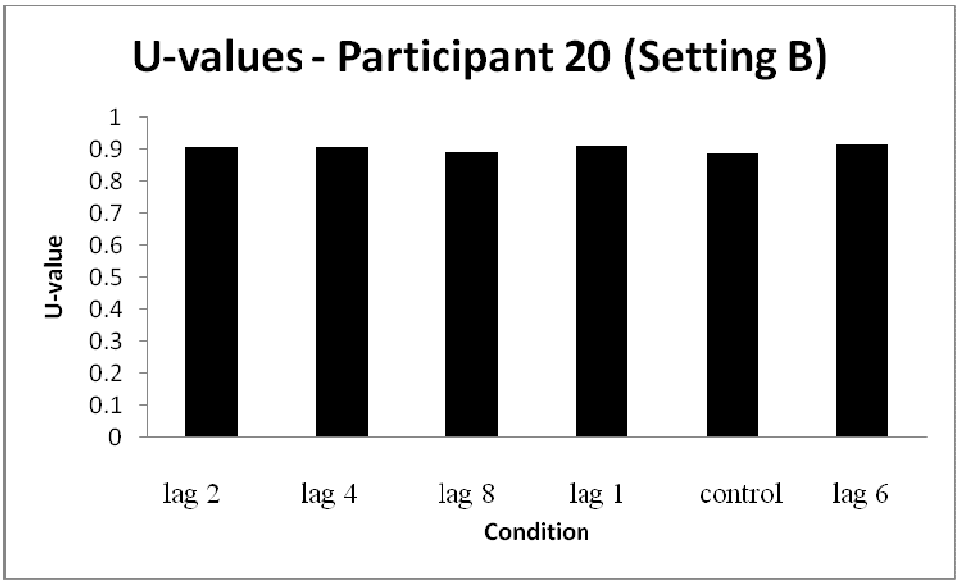


*Figure 4.26:* U-value scores for Participant 20 across the six conditions.

Following analysis of the data from Group 2 and, to a lesser extent Group 1, it is clear that some participants were sensitive to the contingencies in place in each condition and their U-values changed according to the changing lag criterion. However, several of the participants demonstrated little change in their U-values across the conditions (two in Group 1 and four in Group 2). As there is a dearth of research on operant variability among neurotypical children, we sought to examine some of the findings from operant variability research using animal performances to explain the current findings.

Figure 4.27 (reproduced, with permission, from Abreu-Rodrigues et al., 2005) represents U-values obtained by four pigeons across lag 5, lag 1 and lag 10 schedules. The closed bars represent these schedules while the open bars represent the U-values when the pigeons chose an alternative schedule under these conditions. The closed bars are of interest here. Pigeons P10, P20 and P30 responded as would be expected, with the lowest U-value in the lag 1 condition and the highest U-value in the lag 10 condition. However, Pigeon P40 responded in a similar way to some of

the participants in the current study (particularly Participants 5, 6, 11, 14, 15 and 20 ) as the U-value was approximately similar across all conditions, irrespective of the lag schedule in place.

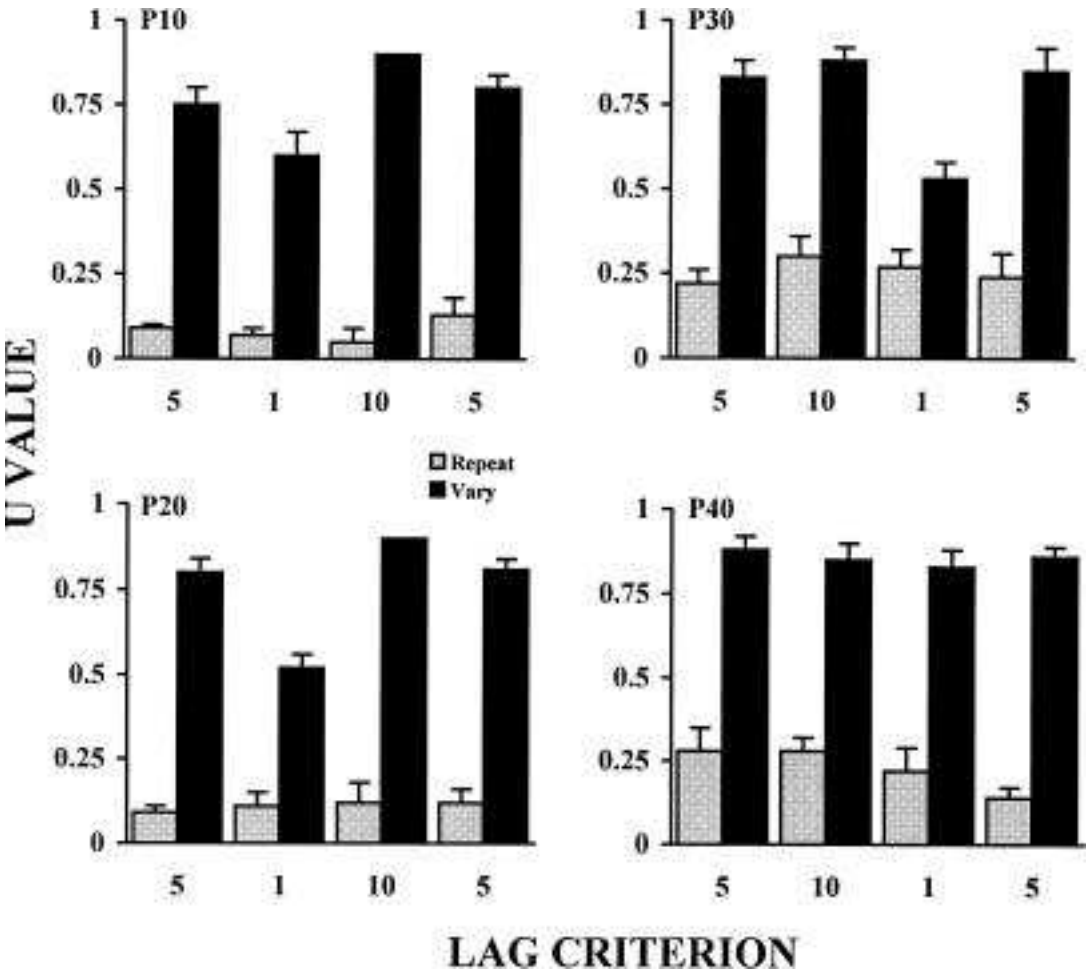


*Figure 4.27:* U-values of four pigeons in REPEAT (open bars) and VARY (closed bars) under different lag schedules (reproduced with permission from Abreu-Rodrigues et al., 2005).

### 4.4. General Discussion

The current study revealed a number of interesting findings. Firstly, the significant between groups effect provides further evidence for the difference in response varibility in children with ASD compared to their neurotypical peers. Group 1 achieved consistently lower U-values across the study conditions than Group 2, indicating lower variability in responding. This reflects the findings of the previous two studies but the current study developed the procedure employed in the previous one in order to identify remediation strategies for this deficit.

Data from this study clearly show that lag schedules of reinforcement can increase response variability. Importantly, they also show that increasing the lag value results in higher variability. The current study indicates that, when using lag schedules with human participants, there is a ceiling, beyond which extinction of variability can occur. This is likely due to the magnitude of reinforcement decreasing while response effort increases. Findings from the current study indicate that a lag 6 may be the optimum schedule for high levels of variability in human participants with and without ASD.

However, there were six participants of the 20 in this study that showed little or no sensitivity to the changing lag criterion. These participants demonstrated relatively high variability when reinforcement was not contingent on it (during the control condition) and this may be relevant in explaining their apparent resistance to the altering lag criterion. On examination of the animal research in this area, Abreu-Rodrigues et al. (2005) reported that one of their four pigeon subjects demonstrated high and stable levels of variability across lag 1, lag 5 and lag 10 conditions. This mirrored the two participants of Group 1 and the four participants of Group 2 who demonstrated this pattern of high and stable variability in responding. Further

research is required to identify which participants are most and least responsive to lag schedules and to attempt to indentify the participant characteristics related to both outcomes.

A preliminary investigation of the relationship between severity of autism sympoms and response variability was conducted as part of this study. There were some indicators that an inverse relationship between these two variables may occur but further investigation is warranted. Future research with a large number of participants may yield useful information regarding the effect of autism severity on the level of response variability in this population.

The additional efforts made to design a bright and engaging interface for this study (relative to the previous study) were extremely worthwhile as the participants were enthusiastic about “playing the game”. It was particularly reinforcing for them when the characters progressed across the levels and this contributed to their overall willingness to engage with the computer program which was the “front” for the lag schedules and subsequent data collection. The positive outcomes of this study are likely to be related to the enjoyment the participants had when playing the game and their motivation to progress through it by fulfilling the various lag criteria.

Considering the prevalence of low variability in responding observed throughout this thesis and the associated RRBs identified in Chapter 2, the importance of identifying procedures for ameliorating invariable response patterns cannot be underestimated. The current study highlights an important avenue for further research as it has demonstrated that the use of lag schedules of reinforcement can yield encouraging outcomes for individuals with ASD who demonstrate low variability in responding.

The current chapter investigated the use of lag schedules using an automated program. Chapter 5 will expand on the current findings by utilising lag schedules to increase variability in a naturalistic setting using play actions with toys. Chapter 5 contains two experiments that aim to identify optimum environmental conditions for promoting variability in play among children with ASD.

# Chapter 5

## Study 4: Experiment 1

## Increasing Variability in Actions with Toys for Five Children with a Diagnosis of Autism Spectrum Disorder Using Lag Schedules of Reinforcement

## 5.1. Introduction

The previous study demonstrated that lag schedules of reinforcement successfully increased variability in responding on a computer-based task. Identifying optimal environmental conditions for increasing variability is an important goal as this may serve as a behavioural cusp for children with ASD, facilitating exploratory and sampling behaviour as well as more typical language and play skills. The current study aimed to investigate the effect of lag schedules on the variability of actions with toys by five children with a diagnosis of ASD. In this way, the procedures derived from experimental work may be used to good effect to target socially significant behaviours, in this case, increased variability in play behaviours. Lag schedules involve the delivery of reinforcement if a response or sequence of responses differs from a pre-determined number of previous responses (Lee, Sturmey & Fields, 2007; Neuringer, 2002). Although the utility of lag schedules is increasingly being investigated with the clinical population, particularly in the area of verbal behaviour (e.g., Susa & Schlinger, 2012), there is a gap in the research literature as the usefulness of lag schedules for increasing variability in functional or symbolic play has not been subject to extensive investigation to date.

Infants typically engage in functional play, which involves using an object or toy as its function indicates, e.g., pushing a toy car along the table saying “broom broom” (Ungerer & Sigman, 1981). Symbolic play is more complex and it usually begins at around 20 months. An example of symbolic play is using an object as if it is something else, e.g., pretending a brick is a bar of soap (Libby, Powell, Messer & Jordan, 1998).

Previous work on 1ag schedules has been restricted to increasing variability in different aspects of block building such as the colour and shape of blocks used

(Napolitano, Smith, Zarcone, Doodkin & McAdam, 2010) and the type of activities selected (Cammileri & Hanley, 2005). Napolitano et al. (2010) used a lag 1 schedule to increase variability in either the colour or form of block structures created by children with ASD. For example, in the colour condition participants received reinforcement if the colour of the block used was different to the colour of the previous one, i.e., placing a red block on a blue block.

Cammileri and Hanley (2005) also utilised a lag 1 schedule whereby participants were required to select an activity that differed from the previous one in order to access reinforcement. As both of these studies utilised lag 1 schedules, participants could alternate between two responses and still access reinforcement each time. However, alternating repeatedly between two responses can be considered to be highly repetitive behaviour, particularly in light of the variable nature of typical play. Research investigating the effects of increasing the lag criterion in studies of this type is warranted.

The ability to behave variably is important for many aspects of successful functioning, such as sampling new ways to reach a goal or generating novel combinations of images or concepts when engaging in artistic or scientific work (Neuringer, 2002). Spinka, Newberry and Bekoff (2001) report that play is ubiquitous in almost all mammals. They propose that the function of play among mammals is “training for the unexpected” (p. 143). During play, animals learn how to deal with the element of surprise that occurs when predators are near. Additionally, the authors observed that animals who change habitat frequently will demonstrate more play behaviours. This suggests that the variability training received through play allows the animals to adapt quickly to changing environments. Similarly, Wenner (2009) described variability as an important aspect of free play in

human children as it fosters creativity and influences social, emotional and cognitive development in later life.

Lee, Sturmey and Fields (2007) reported that that the identification of conditions that influence response variability could lead to procedures that could reduce stereotypy in the clinical population and also increase creative play, among other socially significant behaviours. The aim of the current study was to identify if lag reinforcement schedules could form the basis of such a procedure.

The deficit in play skills in children with ASD is perhaps most noticeable during free play with toys. At these times children with ASD frequently engage in repetitive actions with toys and they often do not appear to emit any meaningful imaginative play (Atlas, 1990). Lewis and Boucher (1995) discuss the lack of spontaneous pretend play in ASD in terms of impairment in generativity. They note that such children can follow instructions during play, but they have an inability to generate their own ideas for play actions. They also describe the play of children with ASD as “repetitious” (p. 118). Charmon and Baron-Cohen (1997) found that children and adolescents with ASD could produce functional play actions and object substitutions when prompted, but they produced fewer novel play acts - with and without props - than controls.

This in turn can affect how these children integrate with their peers (Grossman, Carter & Volkmar, 1997). By increasing variability in the toy play of children with ASD, it is likely that this may serve as a behavioural cusp, allowing for more meaningful play interactions with their neurotypical peers and siblings. As the children emit a wider variety of actions, they begin to access more contingencies in their environments. For example, a sibling will play for longer and be more animated and reinforcing with a child who does lots of different actions with a toy

car, such as driving it over around the floor; flying it through the air; opening and closing the doors and putting people inside, than with a child who repetitively spins the wheels of the car. The play skills will be further improved by accessing this social reinforcement and the child will have opportunities to access increasing amounts of such reinforcement.

It has been suggested by Sutton Smith (1975) that play serves as variability training for later life as it fosters creativity and flexibility that is likely to benefit the individual throughout their lifespan. Therefore, in children who have low variability in their play, it is important to investigate methods for teaching this skill.

Children with autism spectrum disorder commonly demonstrate less sampling behaviour than their peers which leads to an inability to maximise reinforcement in their environment (Mullins and Rincover, 1985). This is likely to be related to impairment in exploratory behaviour among this population. Roeyers and van Berkelaer-Onnes (1994) described the deficit in pretend play commonly observed in children with ASD as a result of these individuals lacking a sense of curiosity and exploratory behaviour in the presence of new toys or situations.

A varied repertoire of actions with toys during play may allow for the selection and shaping of a greater number of diverse actions through contact with naturally occurring contingencies. Consequently, when a recently acquired response or responses comes under such control, other functionally equivalent responses are likely to contact those contingencies, thus increasing the number and variety of responses in a child's repertoire (Shahan & Chase, 2002). The impact of this is that increasing variability in play actions in young children with ASD can have long-term positive benefits as this variability increases as the repertoire expands. This, in turn, results in greater access to reinforcement and improved social integration.

The current study aimed to investigate whether lag schedules could be used to increase variability of actions with toys during free play in applied settings. This has important implications for intervention with this population, for whom meaningful play skills have historically been difficult to teach (Stahmer, 1995).

Furthermore, it was expected that the success of the intervention would be related to specific participant characteristics. In order to make some comment on this, an ADOS assessment was conducted with each participant in order to identify their different profiles and analyse the relative success of the intervention across the different participants. In addition, the Repetitive Behaviour Scale – Revised (Bodfish, Symons, Parker & Lewis, 2000) was administered to determine the prevalence and type of restricted repetitive behaviours (RRBs) demonstrated by each of the participants in order to draw conclusions about the relative success of the intervention in relation to baseline levels of variability in the RBS-R subtypes. Chapter 2 demonstrated that level and type of RRBs was significantly related to variability in this population, with stereotypy and self-injurious behaviour found to be the subtypes most strongly related. By administering the RBS-R to the participants in the current study it was possible to investigate if these types of RRBs would be related to differential outcomes of the procedure.

While many of the prior studies utilising lag schedules with clinical populations were limited to a lag 1 schedule (e.g., Cammilleri & Hanley, 2005; Lee, McComas & Jawor, 2002), the current study aimed to investigate if an increased lag criterion (lag 2, 3 and 4) could be used effectively with this population and in this setting. Finally, it was expected that as variability in play skills increased, rates of motor stereotypy would decrease. To assess this, the frequency of stereotypy was recorded across baseline, intervention and follow-up sessions for all participants. As

noted above, presence of the lower order RRBs, including motor stereotypy were found to be more strongly related to low variability than higher order RRBs such as restricted interests. The presence of a high rate of stereotypy means that a lot of the individual's time is allocated to stereotypy and less time is allocated to important play behaviours. Thus, a reduction of stereotypy is desirable so that a child can allocate a more appropriate amount of time to play, without the interference of stereotypy.

## 5.1. Method

### Participants

Five children aged 4-8 years (mean age: 6.2 years) participated in this study. All participants attended an Applied Behaviour Analysis School where they received 30 hours per week of 1:1 and group tuition. Of the 5 participants, 4 were male and 1 was female (see Table 5.1).

| **Participant Number** | Age | Sex | Number of years of ABA intervention |
|---|---|---|---|
| 1 | 8 | Male | 5 |
| 2 | 8 | Male | 5 |
| 3 | 6 | Female | 3 |
| 4 | 5 | Male | 2 |
| 5 | 4 | Male | 1 |

*Table 5.1:* Participant characteristics.

### Setting

All sessions took place in a quiet room next to the participants' classroom. The participants were familiar with the room, having participated in reverse

integration sessions there on a weekly basis. The toys were laid out on a child-size table and a video camera was positioned on a nearby desk using a tripod. Only the student and the experimenter were present in the room during the sessions and the door was left open at all times. No additional materials were required for data analysis as all sessions were scored retrospectively from video footage.

**Materials**

A range of age-appropriate toys was purchased for the purpose of the study. The toys chosen for this study would commonly be used by children for imaginative play. It is here that children interact with the toys and with each other in a variety of ways. The training toys are listed in Table 5.2 and Figure 5.1. A set of toys to measure generalisation was also purchased. The generalisation toys are listed in Table 5.3 and Figure 5.2. A Samsung video camera and tripod were used to record all sessions. A token mat and 20 tokens were used for reinforcement.

| Toy Set | Components |
|---|---|
| Dinosaur Set | Large dinosaur<br>Cave man figure<br>Cave woman figure<br>Baby figure<br>Dinosaur egg (top and bottom)<br>Baby dinosaur<br>Cave with 3 stacking “rooms”<br>Cave car<br>Swing |
| Police Set | Police station<br>Police man<br>Police woman<br>Police dog<br>Police car<br>Traffic light<br>Stop sign |

*Table 5.2:* Training toys.

*Figure 5.1*: Training toys.

| Toy Set | Components |
|---|---|
| Rocket Set | Large rocket with 3 chambers<br>Space man<br>Space woman<br>Space dog<br>Moon car<br>Alien<br>Crater |

*Table 5.3:* Generalisation toys.

*Figure 5.2*: Generalisation toys.

**Measures**

**Repetitive Behaviour Scale – Revised (RBS-R; Bodfish, Symons, Parker & Lewis, 2000).** The Repetitive Behaviour Scale – Revised) is an informant-based 43 item scale. Items are rated on a 4-point Likert scale from 0, "behaviour does not occur" to 3, "behaviour occurs and is a serious problem".

The original RBS (Bodfish, Symons & Lewis, 1999) consisted of 3 subscales: Stereotypic Behaviour, Self-Injurious Behaviour and Compulsions. A revision of the instrument was provided in 2000 by including items adapted from other validated instruments, including the Autism Diagnostic Interview – Revised (Lord, Rutter & Le Couter, 1994) and the Sameness Questionnaire (Prior & MacMillan, 1973).

For the RBS-R, Bodfish et al. (2000) proposed a 6 factor structure. However, Lam and Aman (2007) carried out an extensive evaluation of the instrument, with almost 3 times as many participants as the original validation. They suggested that a 38 item scale with a 5-factor structure was superior. The 5 factors are Stereotypic Behaviour, Self-Injurious Behaviour, Compulsive Behaviour, Ritualistic/Sameness Behaviour and Restricted Interests. Lam and Aman (2007) also conducted a Chronbach's alpha test to measure internal consistency and found that the values were in the moderately high to high range for all subscales (range =.78 - .91). Interrater reliability was found to be acceptable for all subscales, ranging from .57 to .73.

**Autism Diagnostic Observation Schedule - Generic (ADOS-G, Lord et al., 2000).** The ADOS-G is an assessment of social communication, play skills and the imaginative use of materials for the purpose of diagnosis of ASD. A detailed description of this instrument can be found in Chapter 4 (p.99).

## Procedure

**Generation of target play actions.** Prior to the intervention, two ten-minute observation sessions were conducted with neurotypical children. One was a male, aged 7 and the other was a female aged 5. The purpose of this was to draw up a list of possible actions that could be demonstrated with the toys used during the intervention. These sessions were identical to baseline sessions conducted with the participants with ASD (described below) and were also video-taped in order to accurately note all actions. Once this list was compiled it served as the target play actions to be prompted in the course of the procedure. The list of actions can be found in Appendix A.

**Responses measured.** The following responses were measured across all Baseline, Intervention and Probe sessions.

***Target behaviour.*** Variability in actions demonstrated with toys. This was defined as a discrete action demonstrated with a toy. It could include another toy, e.g., putting a policeman into a car.

***Additional measured behaviours.*** The following behaviours were also measured across all sessions.

*Toy sampling.* This was defined as the child picking up or interacting with a toy for more than 1 second. It was expected that the participants would not sample all of the available toys. However, it was decided that it would be too difficult to accurately apply a concurrent schedule of reinforcement for variability in sampling. In order to create an establishing operation for increased variability in sampling, the individual toys had a limited range of movements (or actions). Thus, to fulfil the higher lag criteria, participants would have to sample different toys in order to emit a variety of actions. It was expected that the establishing operation would lead to

increased diversity in sampling without having to target this dimension of behaviour directly.

*Switching.* This was defined as the movement from one toy to another. This measure was used as additional information as the total number of toys sampled did not fully represent the variability shown in interacting with different toys. For example, it was decided that interacting with each toy in the set once would indicate a high sampling score. However, it is more desirable to switch between toys in a more diverse fashion and to select each toy several times in different orders. To investigate if this dimension of behaviour was increasing as a result of the lag schedules, the number of times participants switched from one toy to another was recorded across all sessions.

*Stereotypy.* This was defined as any repetitive motor action that lasted for more than two seconds. Stereotypy could occur with or without toys. A separate instance of stereotypy was recorded if there was a break of two seconds or more between behaviours. The experimenter was careful not to record repetitive play actions as stereotypy. For example, if a participant rolled a car across the table several times this was not recorded as stereotypy as it is appropriate to repeat this type of action within a play activity. Table 5.4 describes the topography of stereotypy for each participant.

| Participant | Topography |
|---|---|
| 1 | Bang items off surfaces (table, wall and other items) |
| | Tap fingers off items |
| 2 | Flick fingers in front of eyes |
| 3 | Tap fingers off items |
| | Flap hands |
| | Sit on floor and bounce up and down repeatedly |
| 4 | Bang items off table |
| 5 | No stereotypy observed in the course of the study |

*Table 5.4:* Topographies of stereotypical behaviours for each participant.

**Preference assessment.** A multiple stimulus without replacement preference assessment was conducted for each participant to identify effective reinforcers. A pool of five reinforcers was identified for each participant and pictures of these were assembled on a choice board. All participants were proficient in using choice boards as part of their daily routines.

**Baseline sessions.** Baseline data was collected for each child across two sessions. Each baseline session was five minutes in duration. During each session the training toys were laid out on the table and the participant was instructed to "play with the toys". No feedback was provided. If the participant left the area they were redirected using least to most prompting and told "let's play with these toys for a little while". A termination criterion for a session was four instances of walking away from the area. This criterion was maintained across baseline, intervention and post-intervention probes.

**Intervention.** The intervention phase consisted of seven 5-minute sessions with the exception of Participant 5, who completed four 5-minute sessions. The setting was identical to the Baseline phase except that there was a token mat on the

table next to the toys. For the four participants who completed the full seven sessions, the first five sessions utilised the same lag value. The lag value was decided upon following inspection of the baseline data for each participant. The following limits were decided upon prior to the study:

- Participants with very low variability at baseline (fewer than 10 different actions with toys) began the intervention phase at a lag 1 schedule of reinforcement;
- Participants with low variability (fewer than 15 different actions with toys) began the intervention phase at a lag 2 schedule of reinforcement;
- Participants with moderate variability (up to 20 different actions with toys) began the intervention phase at a lag 3 schedule of reinforcement.

For the remaining two sessions the lag value was increased by one (lag 2 moved to lag 3 etc.). The higher lag value was implemented in order to investigate the effects of increasing the lag value on variability in play actions. Participants 1-4 completed five sessions at the original lag value and two sessions at this lag value + 1. Participant 5 underwent two intervention sessions at the initial lag value and the final two sessions at the higher lag value.

Procedures were similar across the five participants. In the interest of clarity, the general procedure for the intervention phase will be described for Participant 1 only and any variations across participants will be described separately. Participants either completed one or two sessions per day and all data collection was completed within two weeks.

***Participant 1.*** The participant was provided with the following instructions: “will you show me how you can play with these toys?” Participant 1 began on a lag 1 schedule as he showed low variability in toy play during baseline. Each action that

differed from the previous one was reinforced with a token and the experimenter also offered social praise, e.g., “I like the way you did something different”. If the participant demonstrated a different action with the same toy, this met the requirements for reinforcement (lag 1). If he emitted a similar action but with a different toy, this also met the requirement for reinforcement (lag 1) as the different toys required the participant to emit a topographically different response.

Response Interruption and Redirection (RIRD) was implemented following any instance of motor stereotypy. If the participant did not emit any response (excluding stereotypy) for five seconds, the experimenter physically prompted an action using least to most prompting. Prompted responses also obtained reinforcement throughout the procedure.

Each action that met the lag contingency was reinforced with a token and tokens were exchanged for tangible reinforcers on an FR20 schedule. This procedure continued until the five minute session was complete. Following a break where the participant left the intervention area, the second session began at a lag 1 schedule of reinforcement.

Participant 1 completed 5 sessions on a lag 1 schedule and the 6th$^{th}$ and 7th sessions were conducted using a lag 2 schedule. During the lag 2 phase an action was reinforced if it differed from the previous two actions. A post-intervention probe was conducted following completion of the seven intervention sessions. It was identical to baseline conditions and it was five minutes in duration.

***Participant 2.*** The procedures for baseline, intervention and probe sessions were identical to those for Participant 1 except that the lag values were different. Participant 2 began intervention on a lag 2 schedule and this was maintained for Sessions 1 through 5; while Sessions 6 and 7 utilised a lag 3 schedule.

***Participant 3.*** The procedures for Participant 3 were identical to those described for Participant 1 and this participant also began the intervention phase on a lag 1 schedule. Sessions 1 through 5 utilised the lag 1 schedule and a lag 2 schedule was in place for Sessions 6 and 7.

***Participant 4.*** The procedures were followed as above but the following variations were in place for Participant 4. He began the intervention on a lag 3 schedule due to his higher levels of variability at baseline. This was maintained for the first five sessions and a lag 4 schedule was utilised for Sessions 6 and 7.

***Participant 5***. This participant attended school on a part-time basis and thus was not in attendance at school during part of the intervention. Because of this it was only possible to conduct a total of four intervention sessions. The procedures for baseline and probe sessions were identical to the other participants. However, Participant 5 received a lag 3 schedule for Sessions 1 and 2 and a lag 4 schedule for Sessions 3 and 4.

**Generalisation.** In order to assess whether improvements in variability in actions with the training toys would generalise to different toys, two probes were conducted with a set of generalisation toys that were never involved in baseline or intervention phases. The probe at baseline for play actions with generalisation toys was identical to the other baseline sessions except that only the generalisation toys were available. No instruction was given in the presence of the training toys and they were not seen by the participants again until the training sessions were complete. At this point, a probe for generalisation was conducted under conditions identical to the probe with the training toys. It was decided to conduct a baseline as well as a probe with the generalisation stimuli for the purpose of pre- and post-intervention comparison.

**Data Collection and Analysis.** All sessions were scored retrospectively from video footage as it was not possible to implement the lag schedules and record data reliably. Data was recorded for each target behaviour and additional measured behaviours. These were tallied for each five-minute session. Total number of different actions, switches from toy to toy and total number of toys sampled were tallied according to their operational definitions for each participant, along with the frequency of stereotypy for each session. The number of prompts provided by the experimenter in each session was also noted.

**Interrater Reliability.** For each participant, 100% of sessions were scored by a second observer, who coded all target behaviours from the recorded video footage, and an interrater reliability analysis using the Kappa statistic was performed to determine consistency among raters.

## 5.3. Results and Discussion

### Participant Information Revealed by the Screening Tools

The Autism Diagnostic Observation Schedule revealed that four of the five participants received scores that indicated autism and one participant received scores that indicated autism spectrum (see Table 5.5). While the total scores on the Repetitive Behaviour Scale –Revised are similar across the five participants, scores varied according to Behaviour Subtype (see Table 5.6). Participants 1, 2 and 3 scored highly on the stereotypic behaviour subscale and lower on the ritualistic/sameness behaviour subscale. Conversely, Participants 5 and 6 scored highly on the ritualistic/sameness behaviour subscale and lower on the stereotypic

behaviour subscale. This participant information was considered to be of value when interpreting outcomes of the current intervention.

| **Participant Number** | **ADOS Scores** | | | | **Diagnosis** |
|---|---|---|---|---|---|
| | **Module Number** | **Communication** | **Social** | **Total Score** | |
| 1 | 1 | 8 | 11 | 19 | Autism |
| 2 | 1 | 7 | 7 | 14 | Autism |
| 3 | 1 | 5 | 11 | 16 | Autism |
| 4 | 3 | 4 | 9 | 13 | Autism |
| 5 | 3 | 3 | 6 | 9 | Autism Spectrum |

*Table 5.5:* ADOS scores and diagnoses.

| **Participant Number** | **Repetitive Behaviour Scale – Revised (RBS-R) Scores** | | | | | |
|---|---|---|---|---|---|---|
| | **Stereotypic Behaviour Subscale** | **Self-Injurious Behaviour Subscale** | **Compulsive Behaviour Subscale** | **Ritualistic/ Sameness Behaviour Subscale** | **Restricted Interests Subscale** | **Total Score** |
| 1 | 6 | 3 | 1 | 1 | 0 | 11 |
| 2 | 6 | 0 | 0 | 0 | 2 | 8 |
| 3 | 7 | 1 | 0 | 1 | 1 | 10 |
| 4 | 2 | 0 | 1 | 5 | 2 | 10 |
| 5 | 0 | 0 | 1 | 7 | 0 | 8 |

*Table 5.6:* RBS-R Scores by Subscale.

As this was a single case research design, results for each participant will be presented separately.

**Participant 1**

Figure 5.3 shows the number of different actions with toys demonstrated by Participant 1 across baseline, training and probe sessions. It is clear that this participant demonstrated a marked increase in the variability of his actions with toys from baseline to the post-intervention probe. In each baseline session, this participant demonstrated only six different actions with the available toys in the five minutes allotted. Because of the low number of actions emitted at baseline, a lag 1 schedule of reinforcement was implemented. This was maintained for the first five sessions and a lag 2 schedule was used for the remaining two sessions. During the intervention phase, the maximum number of prompts issued to this participant per session was three, indicating that the lag schedule had a positive effect on variability in this participant's play behaviour.

Interestingly, this participant demonstrated a large increase in variability in his actions with toys when the lag 1 schedule was first introduced (See Sessions 1-3) and the increase was maintained throughout the training sessions. No upward trend was observed for variability in actions across the first four sessions. Indeed, in Session 4 a clear decrease in variability occurred. This could possibly be a result of a satiation effect. The toys may not have been as motivating at this stage as the participant had played with them several times. The number of actions increased in Session 5, so a decision was made to progress to a lag 2 schedule of reinforcement for Session 6. Interestingly, there was an increase in the number of different actions emitted across the two sessions when the lag schedule was increased from a lag 1 to a lag 2. This suggests that when using lag schedules it is important to keep increasing the lag requirement in order to maintain motivation.

The average number of different actions at baseline was 6. During the training phase, the average number of independent actions emitted per session was 17. During the post-intervention probe, Participant 1 demonstrated 15 different actions with toys. This was only slightly lower than the average during training when the schedule was in place. This shows a good maintenance effect from the intervention under study and indicates an increase of nine actions relative to baseline.

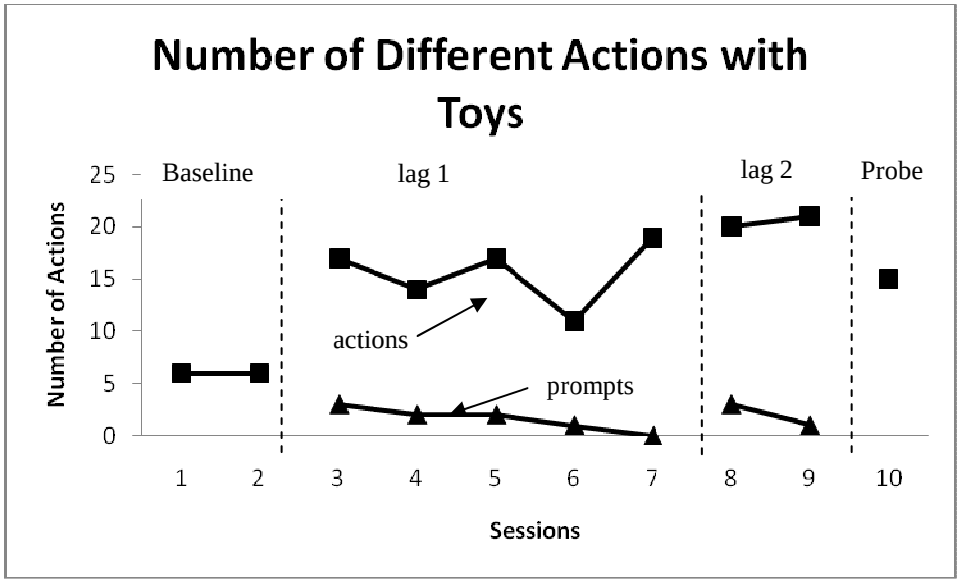


*Figure 5.3:* Number of different actions demonstrated with toys by Participant 1.

The number of different toys sampled in each session by Participant 1 is represented in Figure 5.4. Out of the total of 16 toys available, the participant only sampled five in Baseline 1 and six in Baseline 2. In the training phase, he sampled an average of 10.57 toys. During the probe, he sampled nine toys. Again, this participant showed good maintenance of the increases he exhibited during the training phase.

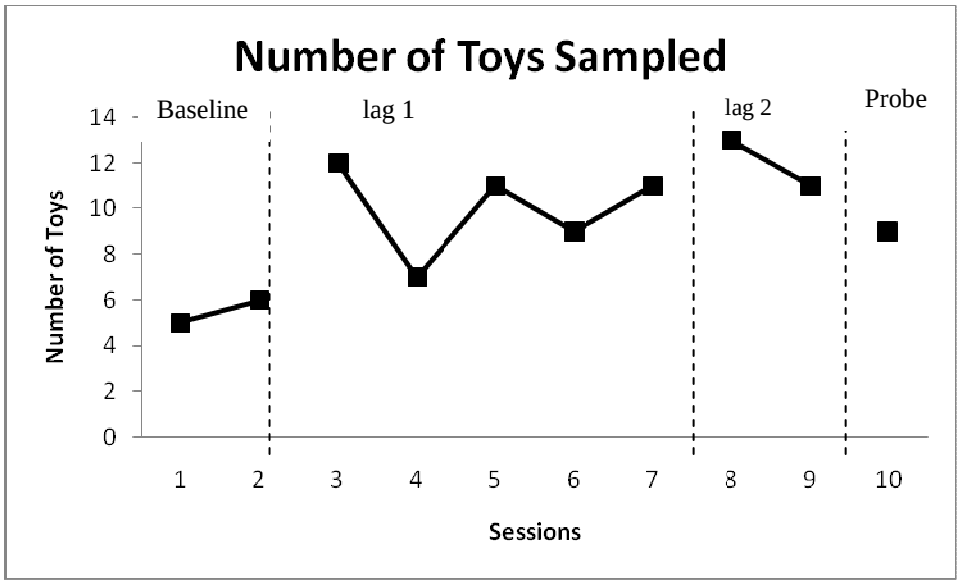


*Figure 5.4*: Number of different toys sampled in each session by Participant 1.

Figure 5.5 shows the number of switches from one toy to another demonstrated by Participant 1. At baseline, he showed 17 and 12 switches respectively. This quickly increased during intervention as the participant was sampling more toys and emitting a wider variety of actions with them. The average number of switches per session during the training phase was 20.71. The number of switches demonstrated during the post-intervention probe was 13, demonstrating a decrease towards baseline on this measure once the lag schedule was removed.

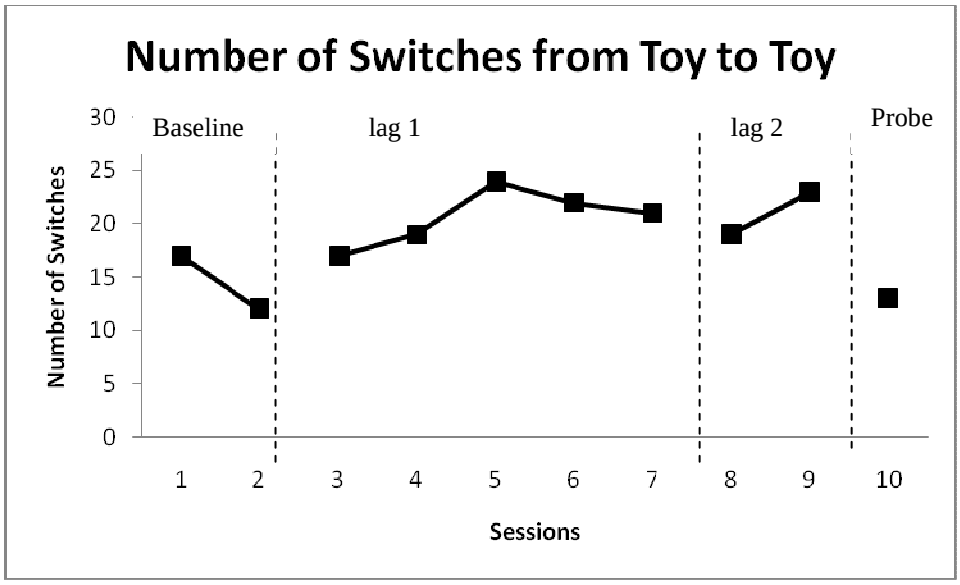


*Figure 5.5*: Number of switches from toy to toy in each session by Participant 1.

Figure 5.6 illustrates the data gathered on the frequency of stereotypy emitted by this participant in the course of the study. At baseline, he demonstrated extremely high rates, with the majority of each five minute session spent engaging in stereotypy. In total, there were 65 instances and 74 instances in Baseline Sessions 1 and 2 respectively (mean = 69.5). In the first session where the lag schedule was introduced, there was a large decrease in stereotypy. This was maintained across the training phase. In fact, the mean number of stereotypic behaviours emitted per session during the training phase was 6.14 (range: 3-9). During the post-intervention probe, 18 instances of stereotypy were recorded. While this was an increase on the average per training session, the large improvement relative to baseline represents a decrease of 74% on the average baseline frequency. It was expected that an increase in variability in play skills would lead to a corresponding decrease in stereotypical behaviours, but the magnitude of the improvement was a welcome outcome of the intervention procedure.

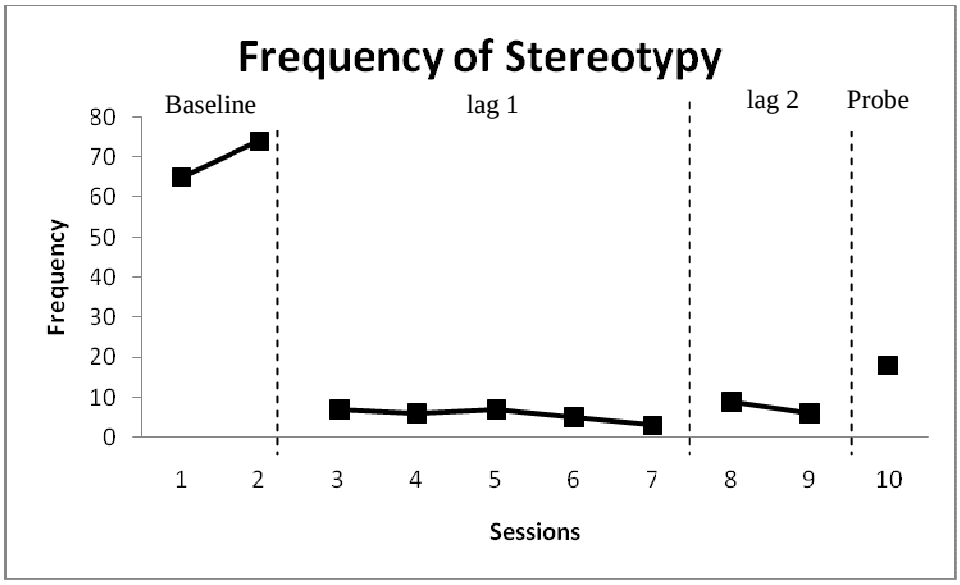


*Figure 5.6*: Frequency of stereotypical behaviours emitted per session by Participant 1.

Figure 5.7 shows the data for number of actions emitted with toys, the number of toys sampled, the number of switches from one toy to another and frequency of stereotypy during baseline and post-intervention probe with the generalisation toys. Participant 1 demonstrated a considerable increase in the number of different actions demonstrated with the toys (from 12 to 20), showing a positive generalisation effect from the training toys to a new set of stimuli. There were small increases in the number of toys sampled and the number of switches from toy to toy, but this may be due to the smaller number of toys available in the generalisation set. This participant sampled six of the seven available toys at baseline and he sampled all seven toys during the probe. The number of switches increased from 23 to 25. The further promising aspect of this participant's data set is the decrease in stereotypy that was demonstrated with the generalisation toys as well as the training toys. This generalisation effect to untrained stimuli is extremely encouraging as the inhibiting effect on stereotypy expected in this study was likely to be under the stimulus control of the toys used in training. Participant 1 demonstrated

29 instances of stereotypy during the baseline with generalisation toys. This decreased to 11 instances following the intervention. These data are most encouraging and point to the efficacy of using lag schedules of reinforcement, both for increasing variability in play skills and for decreasing stereotypical behaviours.

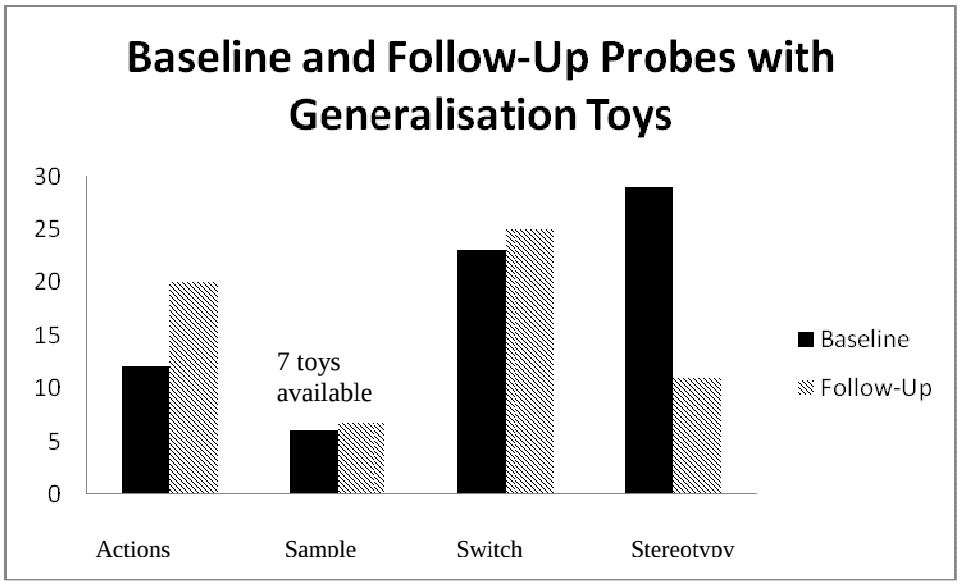


*Figure 5.7*: Results from baseline and post intervention probes with generalisation toys showing number of different actions, number of toys sampled, number of switches from toy to toy and frequency of stereotypy for Participant 1.

**Participant 2**

The number of different actions emitted independently per session by Participant 2 is demonstrated in Figure 5.8. At baseline Sessions 1 and 2 this participant emitted 12 and 13 actions respectively (mean: 12.5). The baseline levels led to a decision to begin the intervention at a lag 2 schedule of reinforcement. The lag 2 schedule was maintained for Sessions 1 through 5 and this was increased to a lag 3 for Sessions 6 and 7. This participant demonstrated some variability in the number of actions emitted during the lag reinforcement sessions. The average number of actions emitted across the intervention sessions was 19.71 (range: 14-25).

During the post-intervention probe, Participant 2 demonstrated 18 different actions, showing good maintenance of increases in the behaviour targeted by the lag schedule. Only two prompts were given to this participant in the course of the intervention and these were in Session 4.

Upon further analysis of the effect of the lag schedule on the variability of actions with toys, it is clear that the highest variability actually occurred during the lag 1 condition in Session 5 (25 different actions). However, the average number of actions during the lag 1 sessions was 19.4. In the lag 2 sessions, 19 and 22 different actions were observed respectively, indicating that increasing the lag requirement by one did not have a significant effect on variability for this participant.

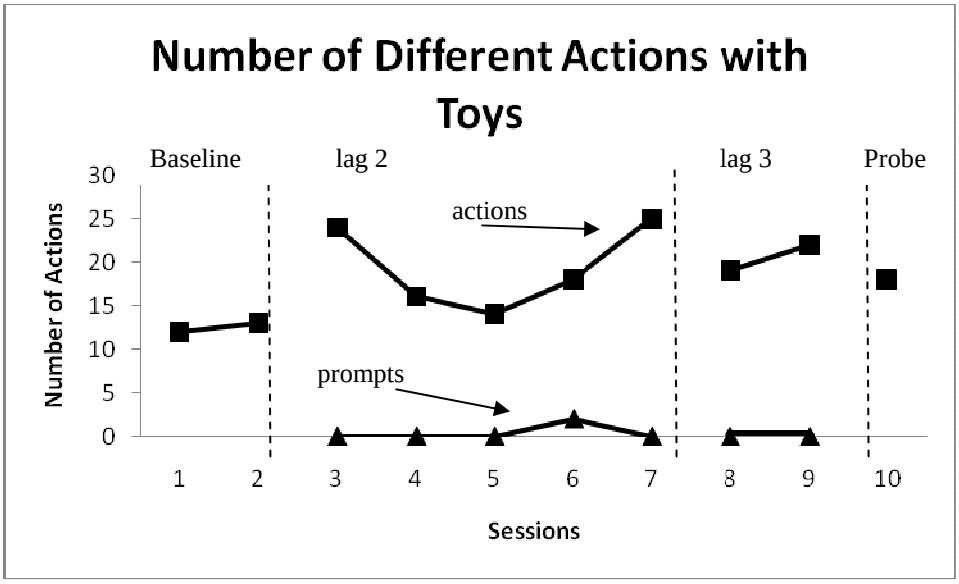


*Figure 5.8*: Number of different actions with toys demonstrated by Participant 2.

Figure 5.9 shows the number of different toys sampled per session by Participant 2. In each of the two baseline conditions this participant sampled 10 of the 16 available toys. During the training phase, he sampled an average of 12.86 toys (range: 9-16). During the post-intervention probe, he sampled 11 toys. Thus, the lag schedule had little effect on the number of toys sampled by this participant.

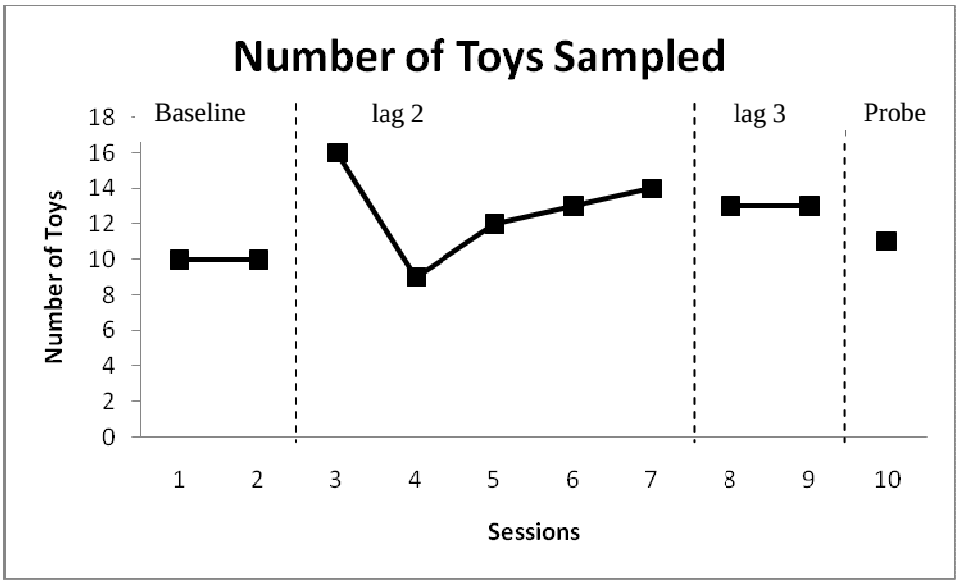


*Figure 5.9*: Number of toys sampled by Participant 2.

The number of switches from one toy to another by Participant 2 is demonstrated by Figure 5.10. During the two baseline sessions, Participant 2 demonstrated 21 and 16 switches respectively (an average of 18.50). In Session 1, this participant showed a peak of 48 switches, suggesting that the initial effect of the lag schedule was to increase switches from toy to toy. Interestingly, in the second session the participant only demonstrated 15 switches, which was lower than baseline levels. During Sessions 3 and 4 he had 20 and 18 switches respectively. It was thought that this was due to a satiation effect following repeated exposure to the toy set. However, the number of switches then increased steadily to 32 switches for each of the final two sessions (lag 2 schedule). The number of switches demonstrated during the post-intervention probe was 20, only a very slight increase from baseline. This suggests that there was no maintenance effect on this measure.

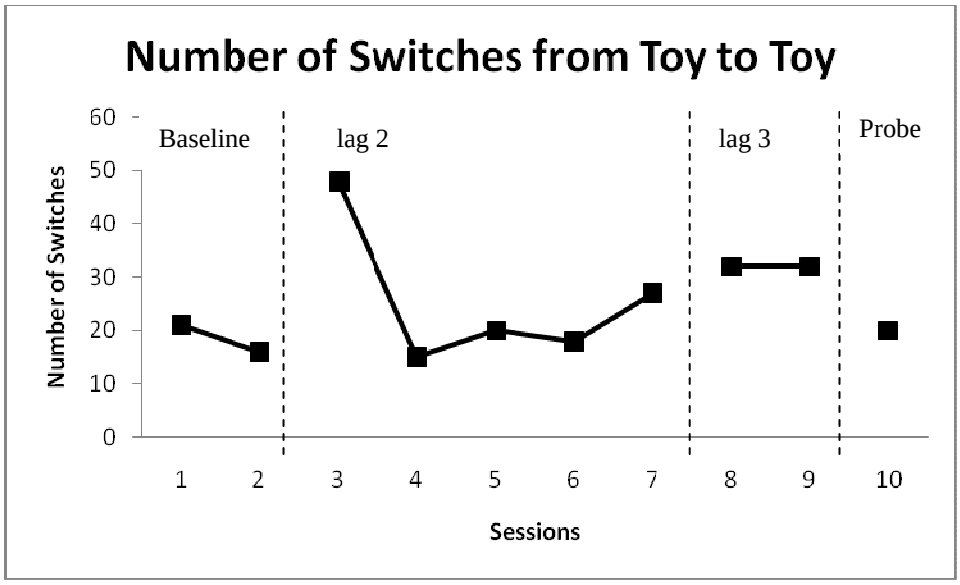


*Figure 5.10:* Number of switches from toy to toy by Participant 2.

Figure 5.11 represents the frequency of stereotypy emitted by Participant 2. During baseline sessions the frequency was two and six (m=4). An average of 0.86 was observed across the intervention sessions (range = 0-3), while there were no instances recorded during the post-intervention probe. While the frequency was low to begin with, it is encouraging to observe the complete absence of stereotypy following the intervention.

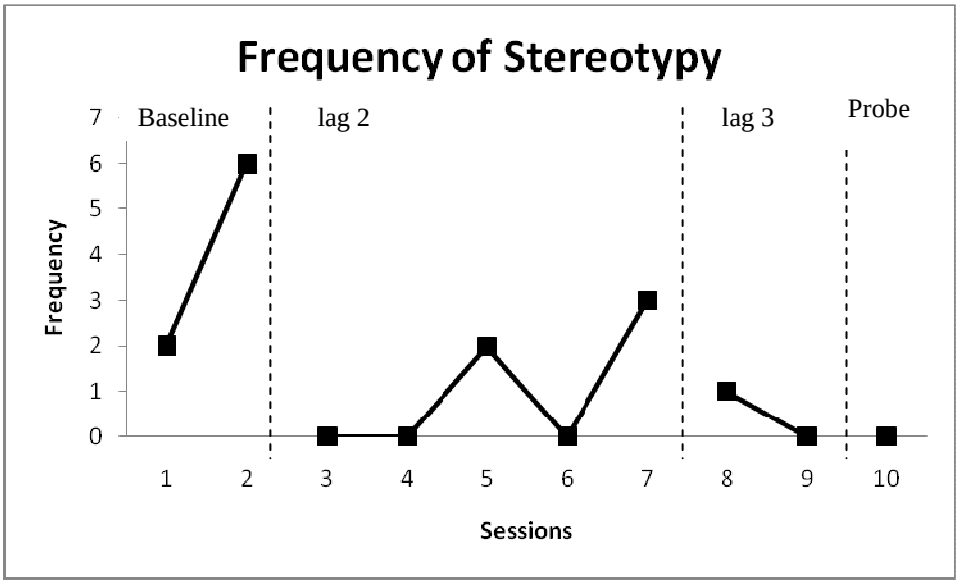


*Figure 5.11*: Frequency of stereotypy emitted by Participant 2.

Figure 5.12 shows the results from baseline and post-intervention probe with the set of generalisation toys across the 4 measures employed in this study. The number of different actions emitted increased from 14 at baseline to 16 following the intervention. The number of toys sampled decreased from all 7 of the available toys at baseline to 5 of the 7 post-intervention. The number of switches from toy to toy increased from 12 at baseline to 14 post-intervention while the frequency of stereotypy was 2 at baseline and 0 at follow-up. These data show modest improvements in variability with the generalisation toys, but, nonetheless, they are encouraging in light of the experimenter's aspiration that lag schedules can show generalisation effects across stimuli and situations.

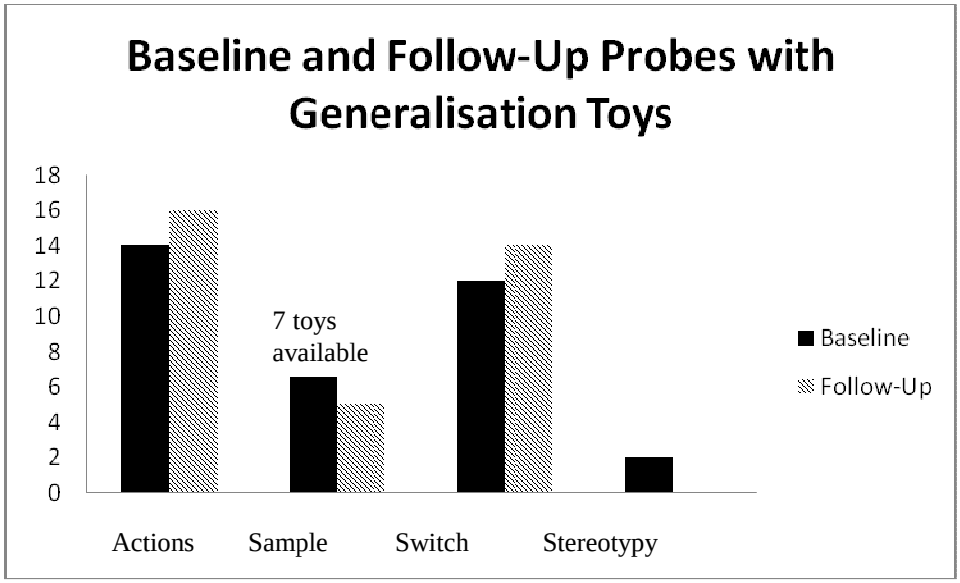


*Figure 5.12*: Results from baseline and post-intervention probes with the generalisation toys showing number of different actions, number of toys sampled, number of switches from toy to toy and frequency of stereotypy for Participant 2.

**Participant 3**

Figure 5.13 shows the number of different actions demonstrated by Participant 3 across baseline, intervention and post-intervention probe. This

participant only completed one baseline session as she reached the termination criterion by refusing to engage with the toys or stay near the toy table during the second baseline session following redirection three times.

From Figure 5.13, it can be seen that this participant only demonstrated one action with the toys during her baseline session. As described previously, intervention began on a lag 1 schedule of reinforcement. The lag 1 schedule was maintained for Sessions 1 to 5 and this was increased to a lag 2 schedule of reinforcement for Sessions 6 and 7.

While the schedule was in place, Participant 3 demonstrated a large increase in the number of independent actions emitted relative to baseline. While there was considerable variation in the number of actions emitted across the lag reinforcement sessions (range: 8-19), there was an overall stability across the intervention phase. The mean number of actions was 15, demonstrating a promising increase from 1 play action at baseline. This indicates a positive effect of the lag schedule. During the post-intervention probe, Participant 3 demonstrated nine different actions with toys. This was considered a favourable result for variability in actions by this participant considering the low variability observed during baseline. The emergence of a maintenance effect on the target behaviour for this participant following seven intervention sessions, the final 2 of which incorporated a lag 2 schedule, was interpreted as a favourable outcome of this exploratory study.

When examining the effect of the lag schedule of variability of actions (as well as in the other measured dimensions) the use of prompting must be considered. Figure 5.13 shows the number of prompts issued per session along with the number of actions emitted independently. The mean number of prompts issued per session during the intervention phase was 7.71 with a range of 1 to 16. The number of

prompts issued decreased rapidly in a linear fashion across the seven sessions to just one and two prompts issued in Sessions 6 and 7 respectively. While the use of prompting may account for some of the variability observed and for the immediate effect of the intervention, the variability in actions was shaped and maintained by the lag schedule. Furthermore, the high level of variability remained as fewer and fewer prompts were required. However, it is clear that a combination of the prompts and the schedule of reinforcement were required in order to effectively increase variability and result in favourable maintenance in the absence of either component.

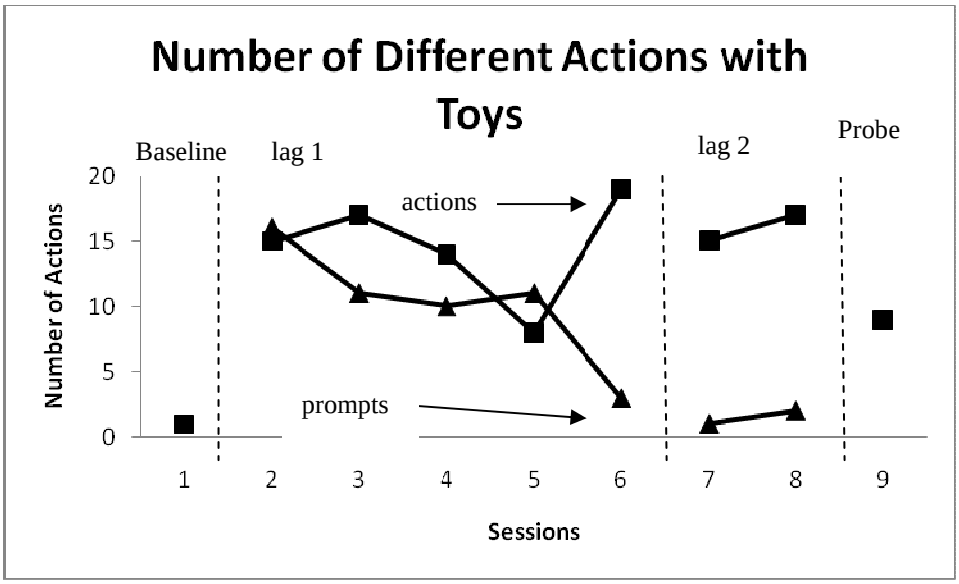


*Figure 5.13*: Number of different actions with toys demonstrated by Participant 3.

The number of toys sampled by Participant 3 across all phases of the study is represented by Figure 5.14. At baseline, she sampled just two of the 16 available toys, indicating very low variability in her sampling behaviour. During the intervention sessions, she sampled a mean of 11.67 toys (range: 11-14). From Figure 5.14, the immediate increase in the number of toys sampled upon the introduction of the lag schedule is clear. These data also show that the variability criterion applied to the target behaviour (variability in actions with toys) acted as an establishing

operation for variability in sampling of toys. During the post-intervention probe, Participant 3 sampled eight of the 16 toys, showing a favourable increase relative to baseline, despite a lower proportion of toys being sampled in the absence of the lag schedule than in its presence.

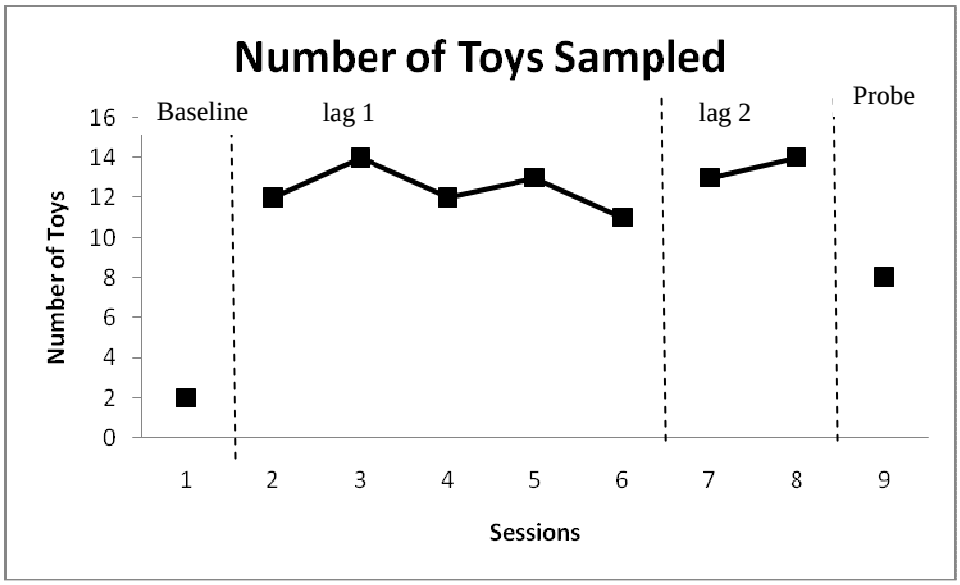


*Figure 5.14:* Number of different toys sampled by Participant 3.

The number of switches from one toy to another that Participant 3 demonstrated during this study is represented in Figure 5.15. During baseline, she emitted only three switches in the five minute session, again highlighting her low variability. As before, this graph highlights the immediate increase in variability once the lag schedule was introduced. During the intervention sessions, this participant demonstrated a mean of 32.29 switches (range: 27-42). During the post-intervention probe, Participant 3 demonstrated 17 switches from toy to toy. This represented a maintenance effect of the intervention on number of switches when the baseline data is considered. From Figure 5.15, it can be noted that the number of switches observed during the probe is approximately 50% of the mean increase relative to baseline. This effect is promising when one considers that this dimension

of behaviour was not specifically targeted by the lag schedule as it was not directly reinforced, but it is considered a side effect of the intervention.

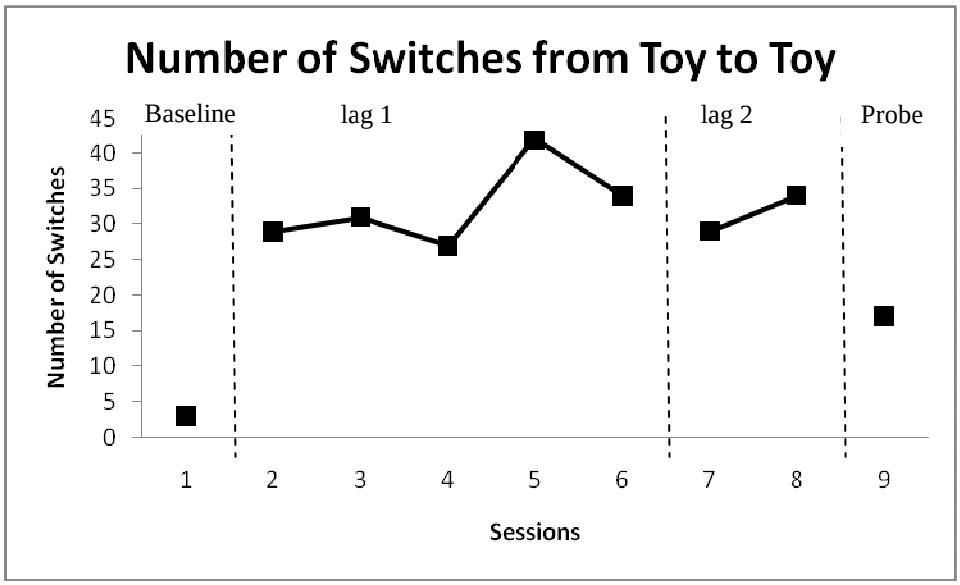


*Figure 5.15*: Number of switches from toy to toy demonstrated by Participant 3.

Figure 5.16 represents the frequency of motor stereotypy emitted by Participant 3 at the baseline, intervention and post-intervention probe phases. During baseline, the frequency was 38 indicating high rates of repetitive behaviour for the five minute baseline. The introduction of the lag schedule resulted in a significant decrease in the frequency of stereotypy. The mean frequency during the training phase was 8.71 (range: 2-19).

During the post intervention probe, the frequency of stereotypy was eight instances in the five-minute period. This represented a 79% reduction in the frequency of stereotypy from baseline to post-intervention probe. This represented an excellent maintenance effect as the frequency during the probe was approximately equal to the average frequency during intervention. Thus, when the lag schedule was removed, even though the levels of variability across the three measured dimensions

decreased, the frequency of stereotypy remained low and stable. These data are extremely encouraging, particularly considering the widely acknowledged difficulties in decreasing stereotypy among individuals with ASD and the impact that such behaviours can have on the individual's social, vocational and academic functioning.

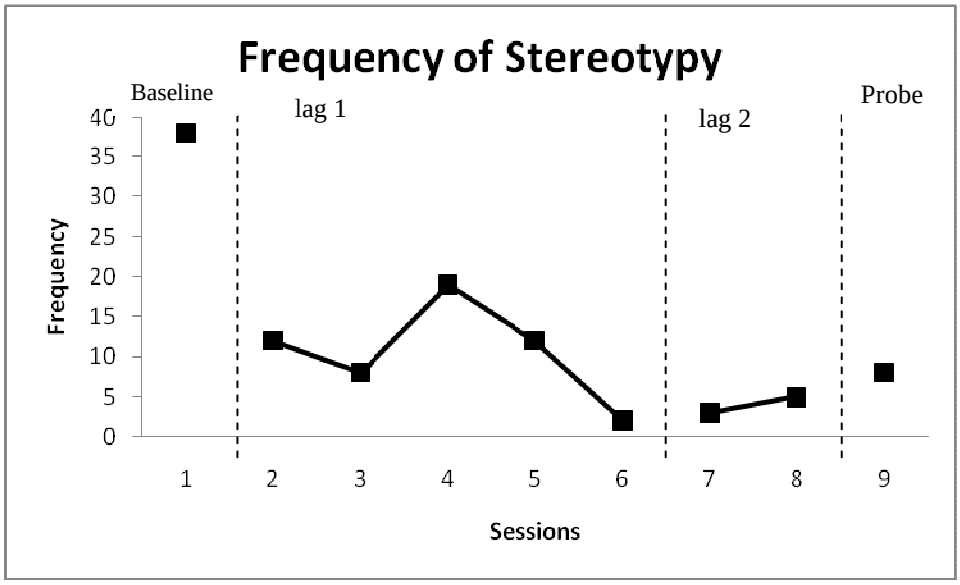


*Figure 5.16*: Frequency of stereotypy emitted by Participant 3.

Figure 5.17 shows the results of baseline and post-intervention probes with the set of generalisation toys. It is clear that for this participant the desired generalisation effect of the lag schedule was not achieved. The number of actions emitted went from nine at baseline to eight at follow-up; the number of toys sampled decreased from five of the seven toys at baseline to four at follow-up; the number of switches from toy to toy decreased from 14 at baseline to 11 at follow-up and the frequency of stereotypy increased from 23 at baseline to 35 at follow-up. These data show trends in the opposite direction to what was desired by the experimenter. There may be several explanations for these data but it is clear that there was no generalisation effect as a result of the intervention.

It should be noted that Participant 3 demonstrated more variability in her actions and sampling with the generalisation stimuli than the training stimuli at baseline. This may have been because she found this set of toys more appealing. Although every effort was made to ensure that the toys were comparable in size, colour/brightness and types of possible actions to be demonstrated it seems that in this case, the participant found the generalisation toys to be more reinforcing as she engaged significantly more with them. However, this effect was not maintained and, in fact, this participant emitted a frequency of stereotypy during the post-intervention probe that approximated the high frequency observed during the baseline with the other stimuli. This may be due to the fact that she required more training sessions in order for the behaviour change observed during the training phase to be maintained and generalised. Further, this participant had high rates of restricted repetitive behaviours as indicated by the RBS-R. It is likely that this influenced the efficacy of the intervention procedure.

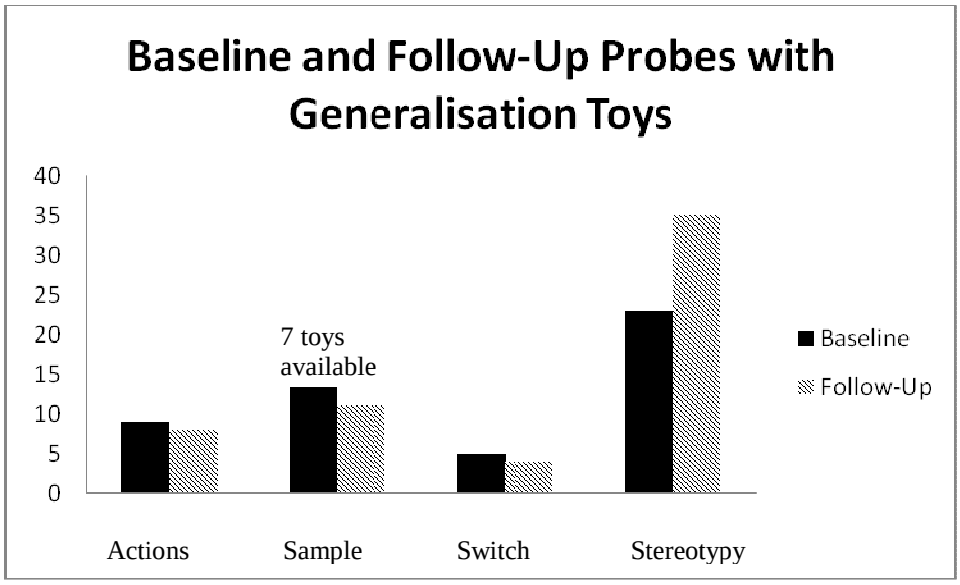


*Figure 5.17*: Results of baseline and post-intervention probe with generalisation toys for Participant 3.

**Participant 4**

The number of different independent actions demonstrated by Participant 4 across baseline, intervention and post-intervention probe sessions is represented by Figure 5.18. From the graph it can be observed that this participant demonstrated higher variability at baseline than the previous three participants. He emitted 16 different actions in each baseline session (compared to six, 13 and two for participants 1, 2 and 3 respectively). Because of this, intervention for this participant began on a lag 3 schedule of reinforcement. The lag 3 schedule was maintained for sessions 1 through 5, while a lag 4 schedule was utilised for sessions 6 and 7. When the lag schedule was implemented, Participant 4 demonstrated an immediate increase to 23 different actions in this 5-minute training session. He emitted an average of 17.29 actions (range: 9-23). During the post-intervention probe, Participant 4 demonstrated 19 different actions with the toys. This was not a large increase on his baseline level of variability but it was favourable in terms of the mean number of actions emitted throughout the intervention phase, showing some maintenance effects.

No prompts were issued to this participant. In hindsight, this may have been a flaw in the research design as it would have been appropriate to interrupt some of the obvious ritualistic behaviour he was emitting with the toys. He quickly learned how to fulfil the criterion of the lag schedule in order to access reinforcement. Thus, while he met the lag requirement, he was observed to line up the toys and to lay them out in particular ways that became quite repetitive and predictable. In short, he learned how to alternate this “ordering behaviour” in order to meet the lag requirement consistently.

Further, in the case of the car, he would roll it along the table in various ways before placing it in the spot he wanted. He would also emit an action with the characters before placing them into their position. He interacted with the toys before placing them in order and was careful to alternate the toys he sampled. Once he had the toys in their order he would select one at a time, emit an action with it and place it back in the same location. He often tweaked his order of placement slightly and in doing so fulfilled the lag criterion. This "ordering behaviour" became more discernible as the intervention progressed and it appeared as if he figured out exactly where he wanted each toy in relation to the rest of the toy set as the intervention progressed.

The behaviour of Participant 4 demonstrated that he responded well to the schedule in place but it also highlighted an issue for careful consideration when implementing this type of intervention with a participant who has a profile like this one. He scored highly (with a score of 5) on the Ritualistic/Sameness Behaviour Subscale of the RBS-R .

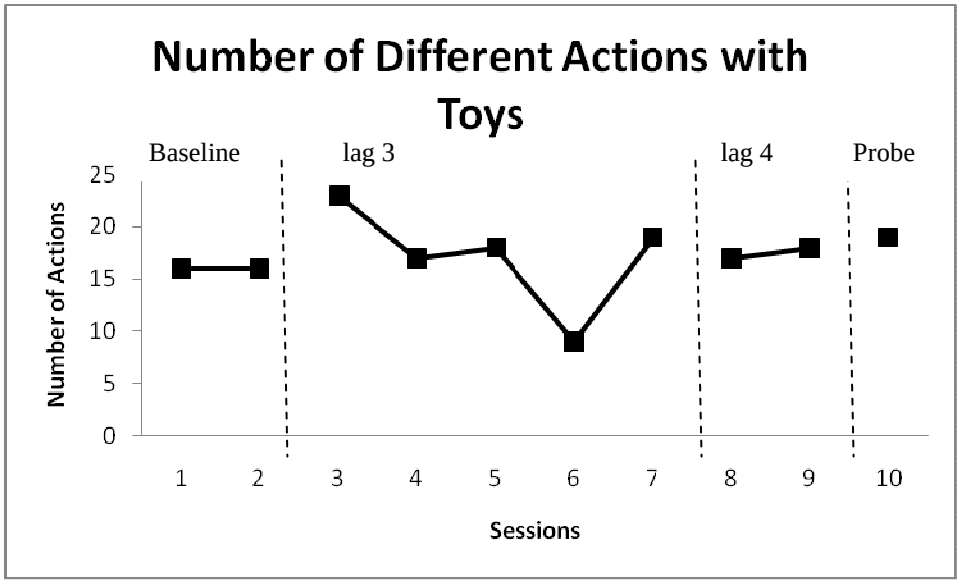


*Figure 5.18:* Number of different actions demonstrated by Participant 4.

The number of different toys sampled per session by Participant 4 is represented by Figure 5.19. During the two baseline sessions he sampled seven and 13 toys out of the available 16. This gave a mean of 10. During the intervention period, Participant 4 sampled a mean of 11.43 (range: 5-14). During the post-intervention probe, he sampled 15 of the 16 toys and showed a promising increase relative to baseline. The mode of the data set was 14, showing that throughout the intervention and during the probe almost all of the available toys were sampled by this participant. This is a most encouraging finding as impaired sampling behaviour in common in children with ASD.

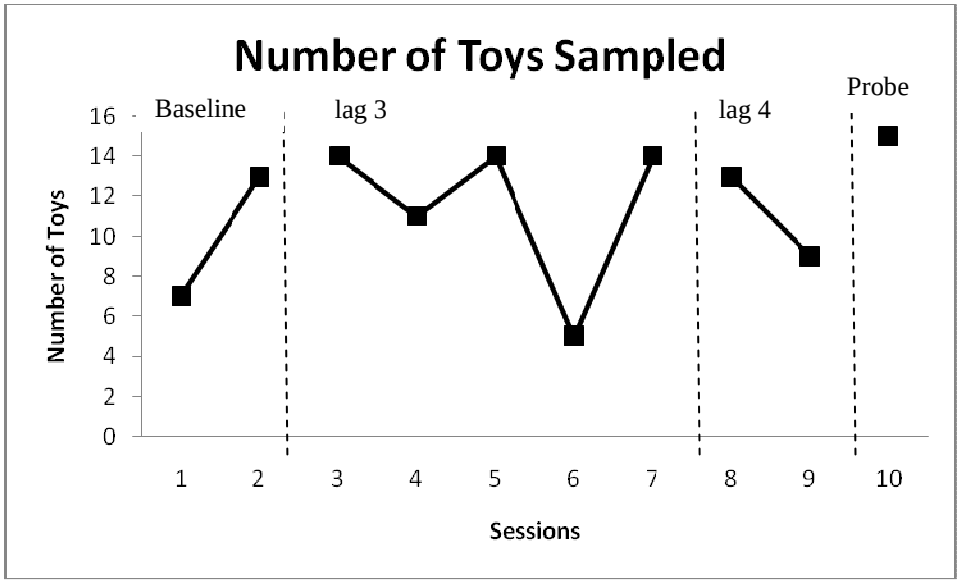


*Figure 5.19:* Number of different toys sampled by Participant 4.

Figure 5.20 shows the number of switches from toy to toy that Participant 4 displayed in the course of this study. From the graph it can be noted that he displayed 12 and 24 switches during the two baseline sessions giving a mean of 18 switches. The discrepancy between the two baselines that was observed in both the sampling dimension and switches dimension for this participant may have been due

to a habituation effect. It is likely that he simply was more used to the toys and the toy table when he completed the second baseline, so he sampled and switched between them at a higher rate. It is interesting to note that the number of actions emitted was exactly the same for the two baselines, even though the sampling behaviour had increased.

During the intervention phase, Participant 4 displayed a mean of 25.57 switches (range: 10-34). The number of switches increased steadily across the intervention phase apart from a dip at Session 4. During the post-intervention probe, he demonstrated 42 switches. This was higher than any figure observed during the intervention phase and represented a major increase on the baseline level.

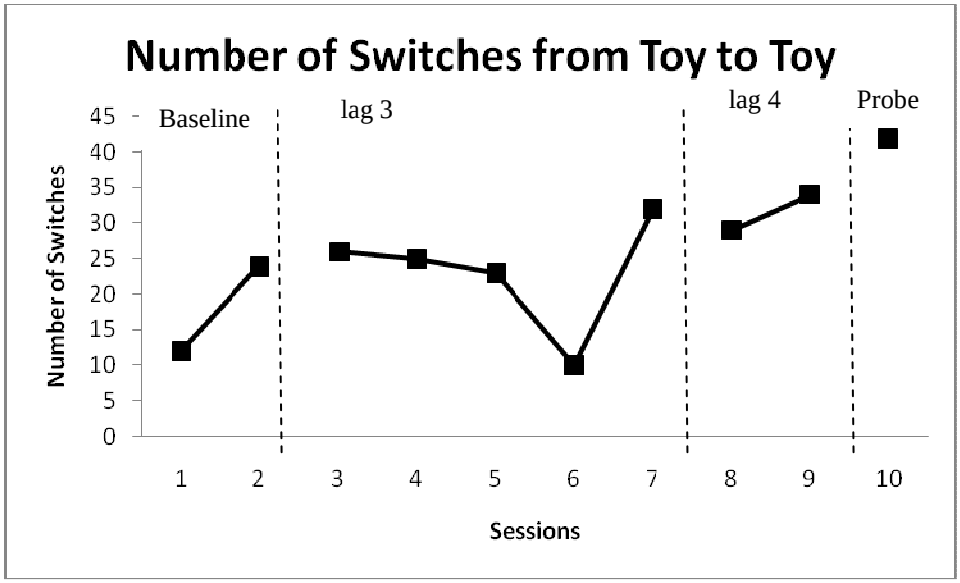


*Figure 5.20*: Number of switches from toy to toy by Participant 4.

Participant 4 did not exhibit any stereotypy at baseline, intervention or probe stages with the training toys.

Figure 5.21 illustrates the results from baseline and follow-up with the generalisation toys. The number of actions emitted with the toys increased from 13 at baseline to 24 at follow-up. This demonstrates an excellent generalisation effect

of the lag schedule. At baseline, Participant 4 sampled six of the seven toys and this increased slightly to seven at follow-up. The number of switches from toy to toy at baseline was 13 and this increased to 25 at follow-up. Again this was a significant increase and further supports the generalisability of increases in variability as a result of lag schedules. Surprisingly, Participant 4 demonstrated some stereotypy during the baseline with the generalisation toys, the frequency of which was 6. However, at follow-up the frequency was 0.

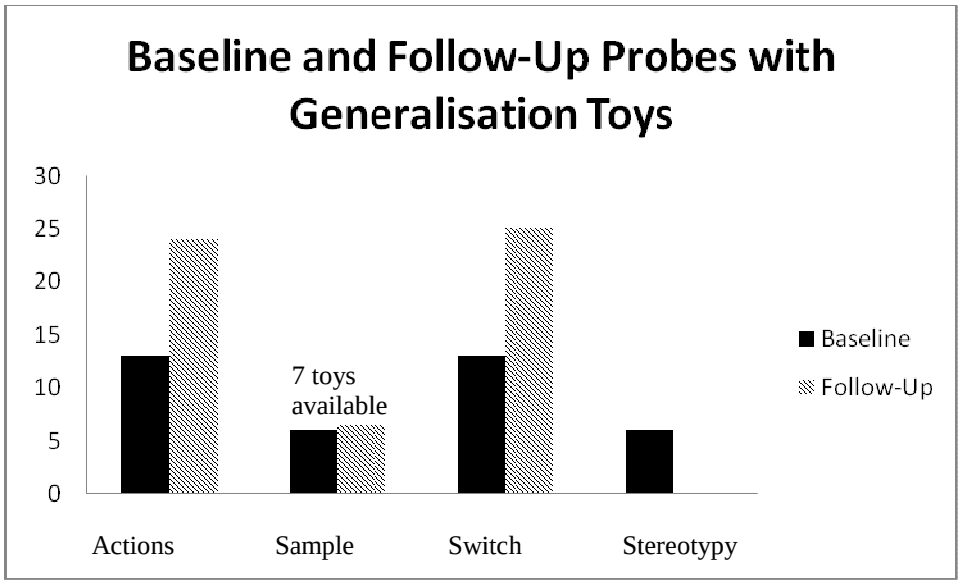


*Figure 5.21:* Results for pre and post intervention probes with generalisation toys for Participant 4.

**Participant 5**

Figure 5.22 illustrates the number of actions that Participant 5 demonstrated with the training toys across baseline, intervention and probe sessions. As previously noted, Participant 5 only underwent four training sessions due to his limited hours at the school where the study was carried out. However, his baseline and probe

sessions, along with all generalisation probes, were identical to those of the other participants.

From the graph, it can be observed that Participant 5 displayed 15 and 19 different actions with the toys during the 2 baseline sessions (mean = 17). Because of this high level of variability at baseline, Participant 5 began the intervention phase with a lag 3 schedule of reinforcement. The lag 3 schedule was maintained for Sessions 1 and 2 and it was increased to a lag 4 schedule for Sessions 3 and 4.

The mean number of actions emitted during the intervention phase was 30.25 (range: 29-32). This represents a significant increase in the variability of actions demonstrated by this participant relative to baseline. Further, the small range indicates that responding was stable across the intervention phase. While this can be considered a positive effect, insofar as the increase caused by the schedule is maintained across the 4 sessions, it also shows that increasing the lag schedule from lag 3 to lag 4 had no effect of the behaviour of this participant. This will be discussed in more detail at a later stage. During the post-intervention probe, Participant 5 demonstrated 35 different actions with the toys, resulting in the highest level of variability observed in the course of the study. Following just four lag reinforcement sessions, he demonstrated variability that represented more than a 100% increase in the number of different actions emitted from baseline to post-intervention probe. This result offers significant support for the use of lag schedules for increasing variability in the play behaviour of children with ASD.

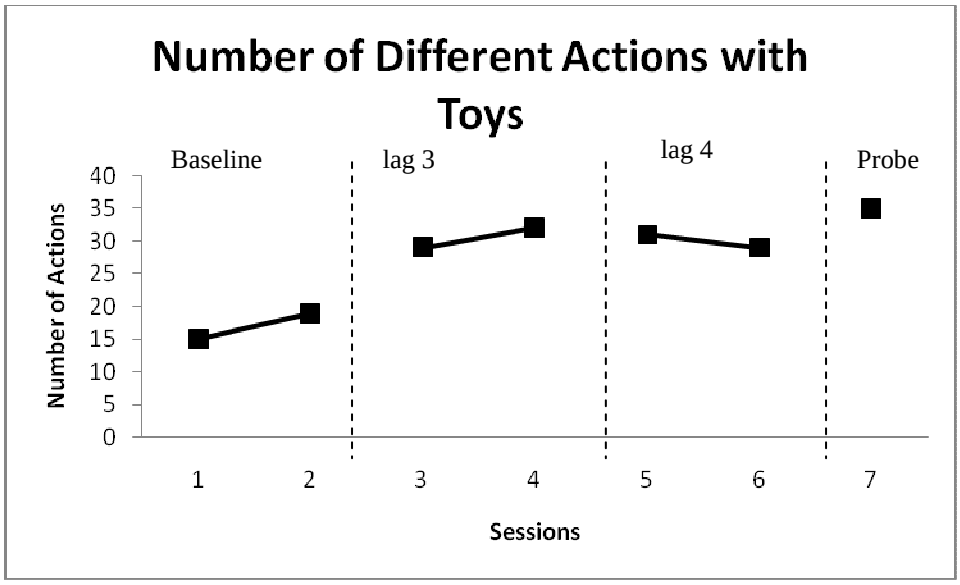


*Figure 5.22:* Number of different actions with toys demonstrated by Participant 5.

Figure 5.23 illustrates the number of toys sampled by Participant 5 in the course of this study. At baseline, he sampled 9 and 12 of the 16 toys in sessions 1 and 2 respectively. During the intervention phase, he sampled a mean of 14 toys (range: 13-16). During the post-intervention probe, he sampled 12 toys showing no increase on his baseline level.

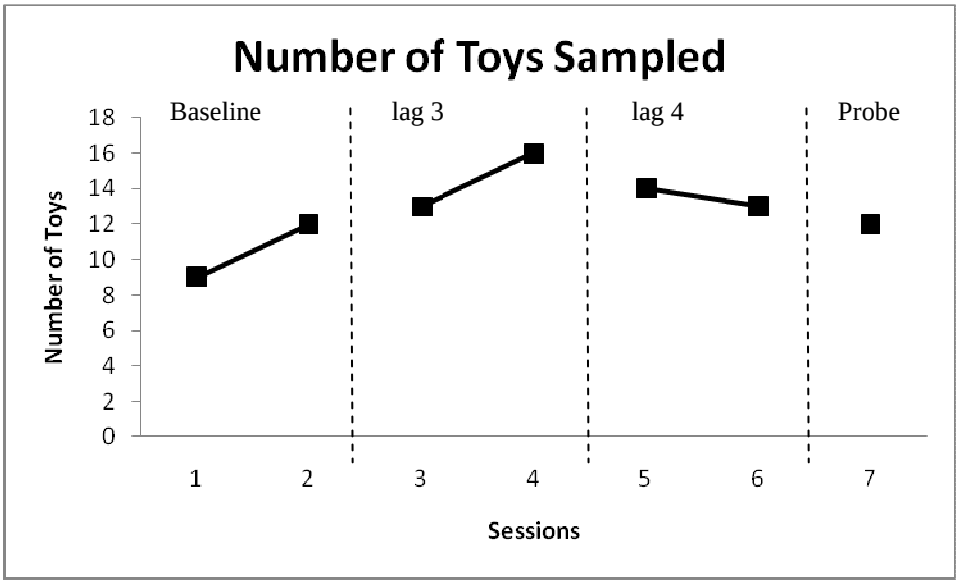


*Figure 5.23:* Number of toys sampled by Participant 5.

Figure 5.24 represents the number of switches from toy to toy displayed by Participant 5. During the two baseline sessions, he showed 34 and 38 switches, resulting in a mean of 36. The mean number of switches observed during the intervention sessions was 36.5 (range: 31-41). The number of switches demonstrated during the post-intervention probe was 37. Again there was no notable increase in variability on this measure.

There was no stereotypy recorded for this participant across the entire study. This is unsurprising on inspection of his RBS-R scores where he scored a 0 on the Stereotypical Behaviour Subscale. Interestingly, like Participant 4, Participant 5 scored highly on the Ritualistic/Sameness Behaviour Subscale. In fact he scored even higher with a score of 7. However, there was no ritualistic behaviour observed from this participant and the research notes actually contain comments about the typical and positive nature of this participant's play, particularly during post-intervention probe sessions where he played independently.

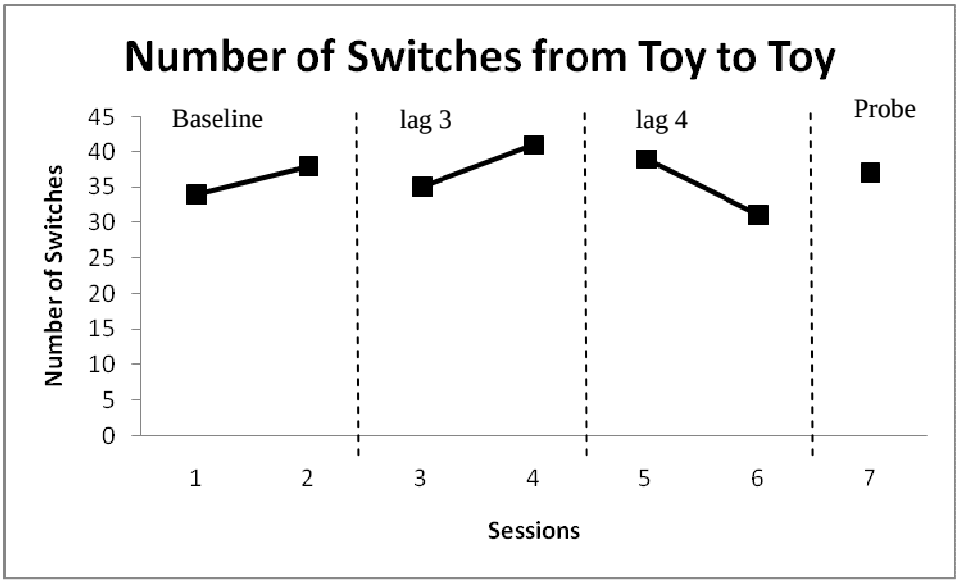


*Figure 5.24:* Number of switches from toy to toy by Participant 5.

Figure 5.25 represents the number of actions, number of toys sampled and number of switches between toys with the generalisation stimuli for Participant 5. It is clear that there was not a large improvement in variability from baseline to probe indicating poor generalisation of the increases observed with the training stimuli. While there was an increase in number of different actions from 24 to 27, the other measures changed in the opposite direction.

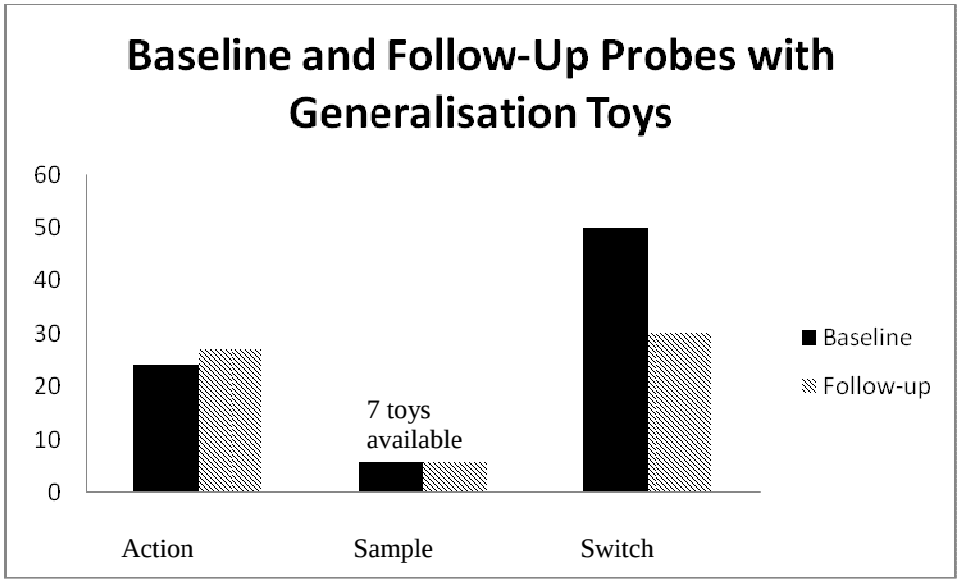


*Figure 5.25:* Results for pre and post intervention probes with generalisation toys for Participant 5.

**Prompts**

All of the scores listed in the above section pertain to independent actions. As noted in the method section, some of the participants required prompting to initially engage with the toys and access the contingencies involved in the lag schedule. Table 5.7 outlines the number of prompts issued to each participant per session across the training phase. Prompts were issued using least to most guidance and were always less than 2 seconds in duration. The criterion for delivering a prompt was no action for 5 seconds or 5 seconds engaging in stereotypy.

| Participant | Session 1 | Session 2 | Session 3 | Session 4 | Session 5 | Session 6 | Session 7 | Session 8 | Total |
|---|---|---|---|---|---|---|---|---|---|
| **1** | 3 | 2 | 2 | 1 | 0 | 3 | 3 | 1 | 15 |
| **2** | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 2 |
| **3** | 16 | 11 | 10 | 11 | 3 | 9 | 1 | 2 | 63 |
| **4** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| **5** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

*Table 5.7:* Number of prompts issued across training sessions for each participant.

**Statistical Analysis**

Table 5.8 displays the mean number of actions, together with standard deviations at baseline, intervention and post-intervention probe for all participants. From Table 5.8, it is evident that there was a large increase in the average number of different actions emitted from baseline to the intervention phase. This represents a promising effect of the lag schedule on variability in actions. There was a very slight decrease in the average number of actions from the training phase (with lag schedule) to the post-intervention probe (no lag schedule), indicating an excellent maintenance effect for the intervention. The larger standard deviation observed in the post-intervention data indicate that there was more variation around the mean in the maintenance phase, but this is likely to be due to the varied participant characteristics at the outset of the study.

| Phase | Mean | Standard Deviation |
|---|---|---|
| Baseline | 10.50 | 6.84 |
| Intervention | 19.85 | 6.05 |
| Post-intervention Probe | 19.20 | 9.65 |

*Table 5.8:* Means and standard deviations for number of different actions in each phase.

A repeated measures ANOVA revealed that the mean number of different actions emitted by the participants in this study differed significantly between baseline, intervention period and post-intervention probe ($F(2,8) = 10.717$, $p = .005$). Mauchleys test revealed that the assumption of sphericity was not violated. The Bonferroni post hoc test revealed that while there was an increase in number of different actions emitted between baseline and post-intervention probe (10.50 ± 6.84 vs. 19.20 ± 9.65), this difference was not significant ($p = .081$). In addition, there was no significant difference revealed between the intervention phase and the post-intervention probe (19.85 ± 6.05 vs. 19.20 ± 9.65) and there was only a slight decrease in variability from the intervention phase to the probe phase ($p = 1.00$).

However, the difference in variability between the baseline phase and the intervention phase (10.50 ± 6.84 vs. 19.85 ± 6.05) was statistically significant ($p = 0.048$), indicating that the intervention had a significant effect on the target behaviour in this study.

**Interrater Reliability**

The interrater reliability for Participant 1 was found to be Kappa = 0.76, $p < 0.0001$. For Participant 2, the interrater reliability was Kappa = 0.695, $p < 0.001$. For Participant 3 it was Kappa = 0.834, $p < 0.001$. For Participant 4 it was Kappa = 0.699, $p < 0.0001$, while for Participant 5 it was Kappa = 0.901, $p < 0.001$. Landis and Koch (1977) described Kappa values of 0.40 to 0.59 as moderate agreement, 0.60 to 0.79 as substantial agreement and values of 0.80 or more as outstanding agreement. All of the Kappa values obtained in the current study were in the range of substantial or outstanding agreement.

### 5.4. General Discussion

As outlined above, the lag schedules implemented in this study had an overall positive effect on variability of actions with toys. All five participants demonstrated a greater number of different actions with the toys at post-intervention probes relative to baseline. Furthermore, all participants with the exception of Participant 4 demonstrated a clear upward trend across the training sessions. The results described above are very promising and indicate that the use of lag schedules is warranted when practitioners seek to increase variability in play behaviours in applied settings. The concurrent decrease in stereotypy from extremely high frequency to low frequency for two of the participants is an additional exciting outcome of the procedure.

All participants but 1 (Participant 3) demonstrated improvements between pre- and post-probes with the set of generalisation toys. The children had no access to these toys for the duration of the training phase but they demonstrated promising improvements in the variability of their actions with them following the lag procedure. This represented a strong test of generalisation as the stimuli were entirely novel and different from the training stimuli in contrast to previous studies where the variations between the training and generalisation stimuli were slight, for example wooden vs. plastic blocks (Napolitano et al., 2010). This is a further encouraging finding as generalisation effects are sometimes difficult to obtain but continue to be a desirable outcome of any intervention. A number of points arise when evaluating the outcomes of the current study as a whole. The decrease in stereotypy observed, as variability in actions increased, indicated that the lag schedule may have served as a DRI (differential reinforcement of incompatible behaviour) for stereotypy. It is desirable in applied settings to teach skills that are

incompatible with stereotypy, not simply reinforcing the absence of the behaviours or “sitting still”. In this way, the current study has succeeded in reaching the goal of habilitation. By increasing variability in play skills the intervention is likely to improve participants’ access to reinforcers (both tangible reinforcement from the toy play itself and social reinforcement arising from increased access to peers due to improved play skills) and also to decrease their access to punishers in the form of reprimands or social stigma associated with high rates of stereotypy.

The variation in the data, particularly those from the post-intervention probe warrants some examination. With such differing participant characteristics at baseline it was expected that the lag procedure would yield differential outcomes according to baseline levels of variability and other participant characteristics. The nature of such characteristics was of great interest to the experimenter.

Firstly, it is interesting to examine the data from Participants 4 and 5. They show that the participants were almost identical in the variability of their actions at baseline (16 for Participant 4 and 17 for Participant 5). However, Participant 4 demonstrated a modest increase across the training phase and emitted 19 different actions during the probe. Conversely, Participant 5 demonstrated a large increase across the training sessions and emitted 35 different actions during his probe. Furthermore, Participant 5 only completed four training sessions while Participant 4 completed seven. Analysis of the profiles of these participants prior to the intervention may hold some clues as to the differing success rates of the intervention for these two boys.

Participant 4 had a total score of 13 on the ADOS-G, giving him a classification of autism while Participant 5 had a score of 9, giving him a classification of autism spectrum. This difference in severity of ASD symptoms may

account for the different outcomes. Moreover, Participant 4 received a score of 10 on the RBS-R while Participant 5 received a score of 5. Interestingly, Participant 5 scored higher on the Ritualistic/Sameness Behaviour Subscale (7 vs 5 respectively). However, Participant 4 scored 2 on the Restricted Interests Subscale and 2 on the Stereotypic Behaviour Subscale while Participant 5 scored 0 on both of these. It is likely that the range of RRBs demonstrated by Participant 4 may explain why the intervention was not as successful in significantly increasing the variability in his actions with toys.

Secondly, if the data set is viewed in its entirety, a similarity emerges between Participants 2 and 4. At baseline they differed in terms of the number of different actions they demonstrated (12.5 vs. 16). This difference was interpreted to be meaningful insofar as they began the intervention on different lag values (2 vs. 3). However, both participants responded in similar ways to the intervention. Both showed an overall increase in variability in the presence of the lag schedule but the data is variable.

Both participants demonstrated a modest increase relative to baseline once the lag schedule was removed, particularly in light of the expectations that these participants would respond very well to the intervention. Participant 2 demonstrated an increase of five actions from baseline to probe; and Participant 4 demonstrated an increase of three actions across this period. During the probe, Participant 2 demonstrated 18 different actions while Participant 4 demonstrated 19, these data show that their scores converged following the intervention. Interestingly, these participants received similar scores from the ADOS-G (Participant 2 scored 14 while Participant 4 scored 13). Severity of autism symptoms may be related to outcomes of a lag intervention.

Likewise, Participants 1 and 3 share many traits, including, similar scores on the ADOS-G (19 for Participant 1 and 16 for Participant 3) and similar profiles on the RBS-R, mainly due to high rates of stereotypy. They also reacted in a similar way to the intervention. They did not emit many different actions at baseline, they made significant gains while the lag schedule was in place and they maintained the improvement during the probe though they showed decreases from the training sessions to the probe session. They also both demonstrated large decreases in stereotypy (>70%).

Participant 5 reacted extremely positively to this intervention. However, he showed the most variability and less severity of autism symptoms at baseline so this was to be expected. Differential severity of ASD symptoms across the participants in this study needs to be considered when discussing its outcomes. Measures such as the ones employed here may also be of practical significance as they can assist practitioners in planning such procedures. Children with greater severity of autism and high rates of stereotypy are likely to require additional intervention strategies such as the ones employed in this study (hierarchy of prompts and RIRD) in order to maximise the benefits of lag schedules of reinforcement.

Participant 4 appeared to find a way to “cheat” the lag schedule and this may indicate a possible flaw that should be considered when using it. As discussed above, Participant 4 fulfilled the requirements of the lag schedule but actually demonstrated some repetitive actions and “ordering behaviour”. The lag schedule may have promoted such behaviour as he was able to access reinforcement in the presence of it. Some ideas for how to address this will be outlined later under directions for future research.

**Advantages of using lag schedules**

The advantages of using lag schedules in applied settings include their efficiency. The positive outcomes of the current study were achieved after only seven five-minute sessions. Indeed, Participant 5 achieved a 100% increase in the number of different actions he emitted after only four sessions. As noted above, the participants in the current study varied widely in terms of their ability and level of variability at baseline but nonetheless, all five participants achieved meaningful improvements in both their variability in toy play and their stereotypy (where relevant).

While lag schedules may seem complex to administer to the naïve observer, this was not the experience of the experimenter in this study. The procedure, particularly when conducted in short bursts, was straight-forward and efficient to implement thereby showing ease of administration in a naturalistic setting.

Lag schedules may hold the answer to the criticism of many ABA programmes, particularly in the area of early intervention, namely that they produce inflexible responding that is difficult to generalise (Powell & Jordan, 1996). By programming for variability from an early stage, children of all abilities can learn to generalise their skills and access naturally occurring reinforcers in their environment more rapidly than usual.

**Future Research**

Future research is warranted to refine our knowledge of the effects of lag schedules on the variability of different repertoires in different populations and to promote their efficacy as an intervention that can be utilised as part of regular ABA programmes.

Further investigation on the participant characteristics that contribute to the relative success of the procedures is warranted, as the data above indicate that these characteristics should influence the planning, implementation and outcomes of such procedures.

In light of Participant 4's ability to "cheat" the lag schedule, research that investigates how best to counteract this would be useful. Research could investigate how to adapt the procedure when the child is engaging in this type of ritualistic behaviour.

Further research into the use of lag schedules for different repertoires such as verbal behaviour and movement in sport may increase the prevalence and popularity of this type of intervention and lead to more researchers and practitioners recognising its value for increasing response variability.

The second experiment of Chapter 5 investigates the efficacy of different lag values by systematically increasing them in a single-subject research design. It aims to identify the most effective schedule for variability in play actions with toys for one participant from Experiment 1, Participant 2.

# Chapter 5

## Experiment 2

## Increasing Variability in Actions with a Child Diagnosed with Autism Spectrum Disorder Using Incremental Lag Schedules of Reinforcement

## 5.5. Introduction

In Experiment 1, evidence was provided to support the use of lag schedules of reinforcement for increasing variability in the toy play of children with autism spectrum disorder (ASD). This is a significant finding and, in the future, it is likely that the use of these schedules with clinical populations will be more commonplace. Lee, Sturmey and Fieds (2007) reported that the identification of conditions that can result in or maintain variability in responding should refine existing practices and aid in the development of new strategies that reduce stereotypy and also increase generative language, problem-solving and creative play. From this initial work, it appears that the application of lag schedules may well be an important step in achieving this objective.

On reviewing the data reported in Experiment 1, some questions arose that warranted further investigation. Experiment 2 aims to answer some of these questions in the pursuit of a greater understanding of the precise mechanisms that underlie the relative successes of lag schedules in producing high levels of variability in responding, in maintaining this variability and in generalisation to new situations and stimuli (in this case novel untrained toys).

To this end, one participant (Participant 2) was selected from Experiment 1 to participate in a comprehensive lag reinforcement procedure. This participant was selected as increased variability in actions with toys was shown by him during the training phase (when the lag schedule was in place) but there was no extra increase when the schedule changed from lag 2 to lag 3. Furthermore, the data were variable across the training sessions. Although he showed more variability during the post-intervention probe relative to baseline, it was not enough to suggest a good maintenance effect. Finally, this participant only demonstrated a slight increase in

the number of different actions emitted with the generalisation toys following the lag schedule procedure.

Napolitano, Smith, Zarcone, Goodkin & McAdam (2010) used a lag 1 schedule to increase variability in the block-building of children with ASD in an ABAB withdrawal design. They noted that, for two of their participants, only a partial return to baseline occurred. This was an interesting finding and while it may be interpreted as a lack of experimental control, it is possible that the increased diversity in responding acquired reinforcing value. In addition, this study investigated generalisation effects to a different set of blocks (wooden vs. plastic). All but one of the six participants demonstrated low variability with the generalisation stimuli following the lag 1 procedure.

The current study will address these issues using an ABAB withdrawal design with generalisation probes across three phases, while systematically increasing the lag requirement in each phase.

## 5.6. Method

### Participant

The participant was a male aged 8 years, 10 months at the time of the study. He had a diagnosis of ASD from an independent Psychologist who made this diagnosis according to the DSM-IV criteria. He had received 5 years of ABA intervention in an ABA school for children with ASD, initially for 30 hours per week but he had been gradually integrated into a mainstream school in the two years prior to the study. At the time of the study he attended his mainstream school for three hours each morning and his ABA school for three hours each afternoon resulting in 15 hours per week in each setting. He had a Special Needs Assistant while in the

mainstream school. In the ABA school he received a mix of 1:1 and group instruction.

The participant had good vocal verbal ability and he typically communicated in 3-4 word sentences. His vocal verbal behaviour mainly consisted of mands and intraverbals and he was reported to have good functional communication. On the ADOS-R (Module 2) his Communication Total was 7 and his Social Interaction score was 7, giving a total of 14, which was above the cut-off of 12 for a classification of autism.

**Setting**

The current experiment was conducted in a quiet room in the participant's ABA school, next to his classroom. Only the participant and experimenter were in the room with the door open. The toys in use at any particular time were laid out on a child-sized table and any additional toys (e.g., those utilised in subsequent phases) were left outside the room.

**Materials**

Four sets of toys were purchased for the purpose of this study. One training set was used for each of the three phases of the intervention and the fourth toy set was used to probe for generalisation. Each training set was comprised of seven component toys. The generalisation set had five component toys. Every effort was made to ensure that the toy sets were age appropriate and that they were standardised across the phases in terms of the number of component toys in each set and the number of possible actions that could be demonstrated with them. The experimenter investigated each of the toys to determine how many concrete and obvious actions

could be completed with each set, (opening and closing doors, placing items in or out of components, driving vehicles, based on the action lists of the previous experiment) and it was determined that 40 typical actions +/- 5 were possible with each training set. Beyond this, additional and less obvious actions, such as causing a horse or a ship to fly, were possible with all toys. Table 5.9 contains a description of the toy sets used in each phase. A Samsung video camera with tripod was used to record all sessions for retrospective scoring.

| **Phase** | **Toy Set** | **Components** |
|---|---|---|
| 1 | Farm Set | Tractor<br>Trailer<br>Fence<br>Cow<br>Horse<br>Pig<br>Farmer |
| 2 | Post Office Set | Post Office<br>Car<br>Postman<br>Postwoman<br>Cat<br>Postbox<br>Letters |
| 3 | Pirate Ship | Pirate Ship<br>Pirate 1<br>Pirate 2<br>Treasure Chest<br>Treasure<br>Sword<br>Armour |
| Generalisation | Ambulance Set | Ambulance<br>Stretcher<br>Paramedic 1<br>Paramedic 2<br>Medical Bag |

*Table 5.9:* Description of toys sets and their component toys used in each phase of the experiment.

*Figure 5.26:* Toy sets used in Experiment 2 (clockwise from top left): Farm Set, Post Office Set, Pirate Ship Set and Ambulance Set.

**Design**

An ABAB withdrawal design with generalisation probes was used in this study. The conditions were identical across each of the three phases (described below) but the lag value increased in each one. The order of the sessions is outlined in Table 5.10.

| Phase | lag value | Conditions | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | lag 1 | A1 | A2 | A3 | B1 | B2 | B3 | gen1 | A4 | A5 | A6 | B4 | B5 | B6 | gen 2 |
| 2 | lag 2 | A1 | A2 | A3 | B1 | B2 | B3 | gen1 | A4 | A5 | A6 | B4 | B5 | B6 | gen 2 |
| 3 | lag 3 | A1 | A2 | A3 | B1 | B2 | B3 | gen1 | A4 | A5 | A6 | B4 | B5 | B6 | gen 2 |

*Table 5.10:* Description of the order of conditions in each phase in the reversal design, where A = Baseline Sessions, B = lag Sessions (lag 1, 2 or 3, depending on phase) and gen = Generalisation Probes.

**Procedure**

The procedure was identical to that of Experiment 1 but the order and number of sessions differed (see Table 5.10). The reinforcers used in this experiment were books and a vibrating ring-pull toy. These were chosen by the participant prior to the experiment. His preferred books were on deprivation at other times so that they were only available to him during the training sessions. All sessions (baseline, training and generalisation) were five minutes in duration. The toys were changed in each phase but each generalisation probe was conducted using the same toy set. All sessions were video-taped and scored retrospectively by two observers.

**Phase 1.** The Farm Set was laid out on a table. As before, the participant was led to the table and given the verbal instruction

*"Here are some lovely toys. Will you play with them for a few minutes?"*

Three baseline sessions were conducted where the participant interacted with the toys in the absence of any attention or interactions from the experimenter. This was followed by three training sessions utilising a lag 1 schedule of reinforcement. As in Experiment 1, the participant received a token if he emitted an action that differed from the previous one. Tokens were exchanged on a FR20 schedule.

Prompts were not used during the training sessions so the lag schedules alone were responsible for any behaviour change. Following three baseline and three training sessions, a generalisation probe was conducted. This was identical to baseline conditions but the specified generalisation toys were used. Next, three baseline sessions, three training sessions (lag 1) and a second generalisation probe were conducted to finish Phase 1.

**Phase 2.** Phase 2 was identical to Phase 1 except the Post Office Set was utilised instead of the Farm Set. Further, a lag 2 schedule was implemented for each of the training sessions.

**Phase 3.** Phase 3 was identical to Phases 1 and 2 but the Pirate Ship Set was used and a lag 3 schedule of reinforcement was in place during the training sessions.

**Data Analysis**

The recorded sessions were scored for the number of different actions emitted, the number of times each action was emitted and the number of toys sampled. It was decided not to record the number of switches from toy to toy as in Experiment 1 as the number of component toys was reduced. The frequency of motor stereotypy was also recorded for the first phase but it was occurring at such a low rate that it was decided that it was not a relevant variable in the current experiment. A second rater scored 69% of the sessions in this experiment and an interrater reliability analysis using the Kappa statistic was performed to determine consistency among the two raters.

## 5.7. Results and Discussion

### Variability in Actions with Toys

The number of different actions demonstrated with toys by the participant across all sessions is represented by Figure 5.27. On visual inspection, there was an increase in the number of different actions demonstrated across the three phases of Experiment 2.

In Phase 1, the participant completed two of the three planned baseline sessions. He refused to engage with the toys during the third baseline session so it was abandoned after 2 minutes and recorded as a refusal. In the two baseline sessions, A1 and A2 he demonstrated 12 and 11 different actions respectively with the toys. When the lag 1 schedule was introduced this increased to 16, 15 and 16 actions with the toys during Sessions B1, B2 and B3 respectively. Next, a generalisation probe was conducted with the specified generalisation toys. The participant demonstrated 8 different actions during this probe.

When the next set of three baseline sessions was implemented the participant demonstrated 13 different actions in each of the three sessions, demonstrating a return to the baseline level of variability in actions. Next, the lag 1 schedule was reinstated for three more sessions and the participant demonstrated 16, 14 and 13 actions in Sessions B4, B5 and B6 respectively. While the reimplementation of the lag schedule resulted in an initial increase in variability in actions, the level of variability began to decrease toward baseline levels by Session B6. It is unclear why this occurred but it is possible that a satiation effect of the toys was responsible. In the second generalisation probe of Phase 1, the participant demonstrated 10 different actions, indicating a slight increase in variability relative to the first probe.

In Phase 2, the participant demonstrated 13, 7 and 11 different actions in the first set of baseline sessions (A1, A2 and A3). When the lag 2 schedule of reinforcement was implemented, this increased to 15, 15 and 16 in Sessions B1, B2 and B3 respectively. Next, a generalisation probe was conducted and the participant demonstrated 16 different actions with the generalisation toys, representing a significant increase in variability in actions when compared to the previous probe.

In the return to baseline conditions with the training toys for Phase 2, the participant demonstrated 16 different actions in the first session (A4), indicating that he was maintaining the increased variability from the reinforcement sessions. However, this decreased to 12 different actions across the next two sessions (A5 and A6). When the lag 2 schedule of reinforcement was reinstated, he demonstrated 12 different actions in the first session (B4) and this increased to 15 different actions in Session B5 and 16 in Session B6, representing a stable level of variability across the lag reinforcement sessions. In the second generalisation probe of Phase 2, the participant demonstrated 14 different actions with the toys, a slight decrease in variability from the previous probe.

In the first set of baseline sessions of Phase 3, the participant demonstrated 14, 12 and 14 different actions (in Sessions A1, A2 and A3 respectively). While this represented a very slight increase in variability in actions relative to previous baseline sessions, it can still be considered a return to near-baseline levels. When the lag 3 schedule was implemented, the number of different actions emitted increased to 17 in B1, 18 in B2 and 22 in B3, representing the highest level of variability in actions in any phase. This indicates that the participant increased his variability in response to the increased lag criterion. Next, the generalisation probe was conducted

and the participant demonstrated 16 actions with the toys, identical to the first probe of Phase 2.

When the baseline conditions were reinstated, the participant initially demonstrated a return to baseline levels of variability, with 12 actions emitted in Session A4. Interestingly, this increased to 22 different actions in Session A5, a repeat of the high level observed in Session B3. In the final baseline session (A6), the participant demonstrated 17 different actions with toys. Although this represented a marked decrease in variability relative to Session A5, it did not represent a return to baseline levels.

In the final set of reinforcement sessions, the participant demonstrated 19 different actions with toys in Sessions B4 and B5 and 18 different actions in Session B6. This represented a slight decrease overall compared to the first reinforcement session in Phase 3. Nonetheless, it indicates that higher variability in actions will be observed when the lag value is increased. In the final generalisation probe, the participant again demonstrated 16 different actions with toys. This is identical to the previous session. Thus, while a clear generalisation effect of variability in actions to novel untrained toys was observed across the phases as a result of the lag schedule procedure, it appears that a ceiling of 16 different actions was reached in this instance. Figure 5.28 represents the outcomes of the generalisation probes in isolation.

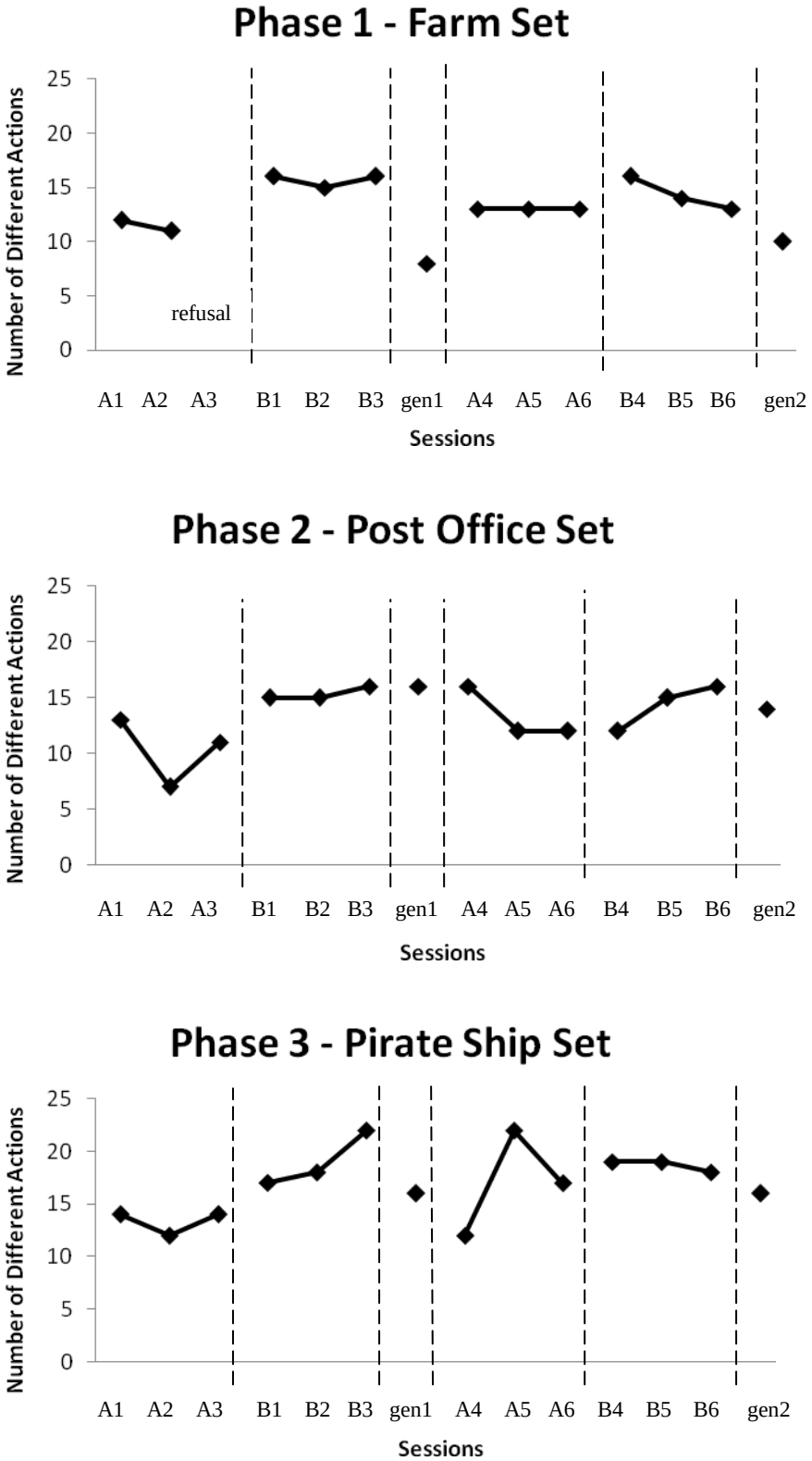


*Figure 5.27:* Number of different actions demonstrated with toys across all sessions (baseline, lag reinforcement and generalisation) for Phase 1, Phase 2 and Phase 3.

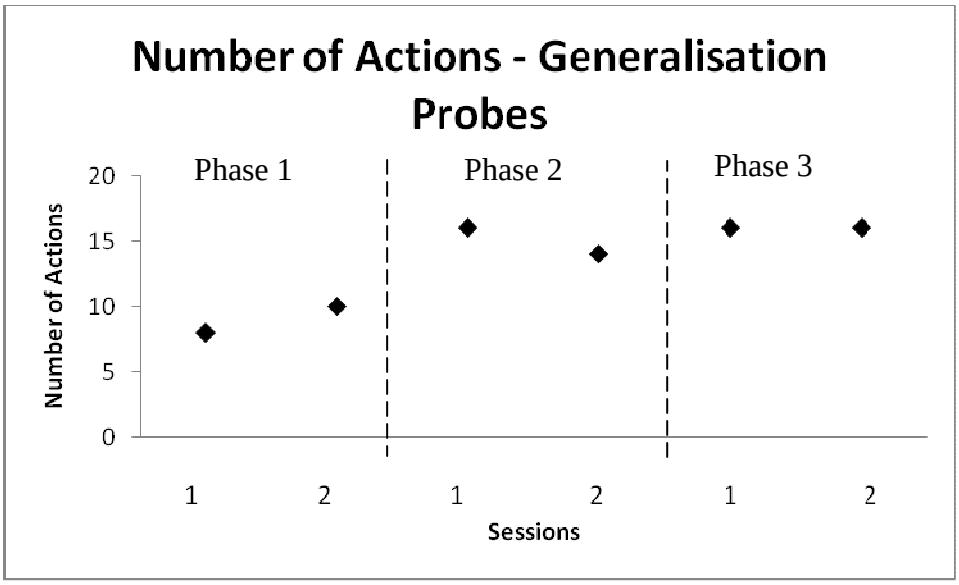


*Figure 5.28:* Number of different actions demonstrated with toys in each of the generalisation sessions of Experiment 2.

**Frequent Actions**

In the course of the study it was observed that the participant tended to repeat many of the actions with toys at a high rate. It was decided to investigate this as it was a variable of interest. For example, if the participant emitted a large number of different actions but then emitted a subset of these many times, the behaviour may be considered repetitive. In order to compile a list of such actions a cut-off was needed to determine what would be considered high frequency. To this end, 20% of sessions (n = 8) were examined to find the mean frequency of actions. There was a total frequency of 513 instances across 118 actions (as actions were counted once for each time they appeared in a session, a single action could be counted up to 12 times), giving a mean of 4.35. Thus, if there were 5 or more instances of the same action within one session this was determined to be a high frequency.

Table 5.11 lists the high frequency actions observed in each phase along with the frequency of each action per session. Note that the frequency of an action is only

listed if it is greater than five in a given session. In Phase 1, it is clear that the participant emitted several of the actions at a very high rate, demonstrating that although the number of different actions increased in the presence of the lag schedule, the high frequency of particular actions can be considered quite repetitive. Indeed, across the entire phase there were 19 instances of an actions being emitted 10 times or more in a 5 minute session.

In Phase 2 there were also many actions at a frequency of 5 or more and these were distributed across the phase. There was only one action that was demonstrated 10 or more times (puts letter in postbox). This action was considered high frequency ($\geq 5$) in 10 of the 12 sessions and, of these, it occurred at a frequency of 10 or more five times.

In Phase 3, a higher level of different actions was observed and there were also considerably fewer high frequency actions in the sessions of this phase. There were only four instances of an action occurring 10 or more times per session in this phase. Across the final seven sessions of Phase 3, there were only one (in Sessions B3, A4, A5, B4) or two (in Sessions A6, B6) high frequency actions per session, with no actions meeting this criterion in Session B5.

| Phase | Toy | Action | A1 | A2 | A3 | B1 | B2 | B3 | A4 | A5 | A6 | B4 | B5 | B6 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Cow | Hop | 10 | 7 | | 10 | 8 | 9 | | 5 | 5 | | | 20 |
| | | Jump | 14 | 5 | | 10 | 6 | 8 | 6 | | | 7 | 5 | 19 |
| | | Lie down | | | | | 6 | 9 | 6 | | | 7 | | 19 |
| | Horse | Hop | 14 | 9 | | 8 | 7 | 5 | 7 | 6 | | | 7 | 15 |
| | | Jump | 14 | 9 | | 10 | 7 | 6 | 8 | 5 | | 6 | 8 | 15 |
| | | Lie down | | 8 | | 8 | 7 | | | 5 | | 8 | 8 | 15 |
| | Pig | Hop | 13 | 8 | | 5 | 5 | 7 | 9 | 7 | 5 | | 7 | 23 |
| | | Jump | 18 | 7 | | | 8 | | 8 | 5 | | 6 | 8 | 24 |
| | | Lie down | | 5 | | | 6 | | 10 | 5 | 5 | 7 | 7 | 23 |
| | | On horse back | | | | | | 5 | | | 6 | | | |
| **Phase** | **Toy** | **Action** | **A1** | **A2** | **A3** | **B1** | **B2** | **B3** | **A4** | **A5** | **A6** | **B4** | **B5** | **B6** |
| 2 | Post Office | Open door | | | | | | 5 | 9 | | | | | |
| | | Close door | | | | | | | 7 | | | | | |
| | | Put item in | | | | 7 | 5 | 7 | 7 | | 5 | 7 | | 6 |
| | | Take item out | | | | 5 | | 6 | 6 | | | | 6 | |
| | Car | Roll straight | 9 | | | 5 | | 5 | | 8 | 6 | 6 | | |
| | | Roll curved | 9 | | | | | | | | 6 | 6 | 5 | |
| | | Put item in | 7 | | | 5 | 6 | | 7 | | | 6 | 7 | |
| | | Item out | | | | 5 | 5 | | | | | 5 | | |
| | Letters | Put in postbox | 8 | 19 | 20 | 15 | 18 | 10 | 6 | 7 | | 9 | 5 | |
| | | Take out | | 7 | 8 | | | | | | | | | |
| | | Place in P.O. | | | | | | 5 | | | | | 5 | |
| **Phase** | **Toy** | **Action** | **A1** | **A2** | **A3** | **B1** | **B2** | **B3** | **A4** | **A5** | **A6** | **B4** | **B5** | **B6** |
| 3 | Ship | Roll straight | 7 | 5 | 5 | 10 | 9 | 6 | | | 6 | 5 | | |
| | | Roll curved | | | | | | | | | | | | |
| | | Press dragon tail | | 5 | | | | | | | | | | |
| | | Open dragon mouth | 8 | 14 | 8 | | 5 | | 10 | | 7 | | | |

| Phase | Toy | Action | A1 | A2 | A3 | B1 | B2 | B3 | A4 | A5 | A6 | B4 | B5 | B6 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Open cabin door | | | | 5 | | | | | | | | 7 |
| | | Close cabin door | | | | 5 | | | | | | | | 12 |
| | | Put item on level 2 | 6 | | | | | | | | | | | |
| | | Lift wings | | | | | | | | 7 | | | | |
| | | Press dragon head | | 9 | 6 | | 5 | | | | | | | |
| | | Press blades/ waves | | | | | | | | | | | | |
| | | Put item on level 3 | | | | | 6 | | | | | | | |
| | Treasure chest | Open chest | | | | | 5 | | | | | | | |
| | | Close chest | | | | | 5 | | | | | | | |

*Table 5.11:* List of high-frequency actions and frequency at which they were observed in each session (when frequency was ≥5).

**Interrater Reliability**

The interrater reliability for the raters was found to be Kappa = 0.879, $p < 0.001$ for Phase 1, Kappa = 0.802, $p < 0.001$ for Phase 2 and Kappa = 0.762, $p < 0.001$ for Phase 3. As mentioned in the previous experiment, Landis and Koch (1977) described Kappa values of 0.80 or more as outstanding agreement. The was surpassed for Phases 1 and 2 while the Kappa for Phase 3 approached this standard and was in the range of substantial agreement.

### 5.8. General Discussion

The outcomes of Experiment 2 further clarify the effects of lag schedules of reinforcement on variability in responding. In the current experiment, the design indicated that the target behaviour of variability in actions with toys was under the control of the lag schedules in place across the different sessions.

The highest level of variability in actions was observed in Phase 3, which utilised a lag 3 schedule, the highest lag value used in this experiment. This is an extremely positive finding as it is the first study to systematically increase the lag value in an applied research study into play behaviours in a clinical setting. This highlights the requirement for many other such studies with larger participant numbers and a variety of clinical populations and target behaviours.

As reported above, in Phase 1 Sessions B5 and B6 (the final reinforcement sessions in Phase 1) there was a decrease in variability in actions with toys. It was observed that this may due to a satiation effect of the toys. Future studies could attempt to remediate this by rotating the toys used. An alternating treatments design using different lag values is likely to be useful for this purpose. Alternatively, toys could be rotated within the same lag schedule. This effect was not observed in Phase 2 and in Phase 3 the effect was so small that it is negligible (a decrease of one action). This suggests that increasing the lag value may also act as a remediation strategy for this effect.

One aim of Experiment 2 was to investigate generalisation effects of lag schedules. The data reported above show that there was a large increase in variability in actions demonstrated with the generalisation toys. Indeed, the number of different actions increased from 8 in the first probe to 16 in three of the four final probes (in Phase 2 and Phase 3). It is unclear whether it was a practice effect from

the number of training sessions or the increased lag requirement that caused this generalisation effect of variability in actions. Further research is required with a greater number of participants to investigate the nature and extent of generalisation effects under different lag conditions. Heldt and Schlinger (2012) reported that a maintenance effect was more likely with a lag 3 schedule than a lag 1 schedule as a lag 3 schedule more closely approximates the contingencies that are naturally occurring in the environment. This may account for the improved generalisation effects observed in the higher lag conditions.

The level of variability in actions with toys was almost identical in the lag reinforcement sessions in Phase 1 and Phase 2 while a clear increase was observed in Phase 3. Further research is required to investigate the implications of this. Perhaps, when utilising lag schedules in applied settings it may be appropriate to progress from a lag 1 schedule to a lag 3 schedule in order to increase variability in a more efficient manner. However, it is important to consider the possibility that it was the cumulative effect of the training sessions across the three phases that resulted in the positive outcome observed in Phase 3.

An analysis of high frequency actions is useful in studies such as this to thoroughly investigate the effects of lag schedules on variability in behaviour. As reported by Lee et al. (2007), lag 1 schedules can result in responding that meets the criteria for reinforcement but it can still be considered repetitive in nature. This is supported by the findings of Experiment 2 as many actions occurred at a high rate (up to a frequency of 24 in one session). It is encouraging to note that increasing the lag requirement remediated this effect as the number of high frequency actions observed by Phase 3 was greatly reduced, indicating that the participant showed more variability in the number of different actions emitted and also, he did not repeat

individual actions at a high frequency, as in Phase 1. Experiments such as this allow researchers and practitioners to further understand the implications of using lag schedules with the clinical population.

Experiment 2 reiterates the finding of Experiment 1, demonstrating that lag schedules are an effective and efficient way to increase variability in toy play in applied settings. While further research is needed to evaluate the optimum conditions under which the desired levels of variability can be achieved across different repertoires, the current outcomes indicate that this is a promising area of study that may have wide-reaching implications for increasing the quality of service-provision for those who demonstrate low variability in their behaviour.

# Chapter 6

# General Discussion

## Operant Variability: An experimental analysis of consequential control over response variation

## 6. 1. Overview of Aims

The current thesis aimed to add to the existing research on operant variability, in particular the prevalence and impact of low variability in individuals with a diagnosis of autism spectrum disorder. Operant variability describes variability in responding, i.e., whether responses are predictable or random-like, as operant behaviour under the control of the consequences that follow it. Numerous studies from the Experimental Analysis of Behaviour (EAB) have provided support for the operant nature of variability (Denney & Neuringer, 1998; Page & Neuringer, 1985; Pryor, Haag & O'Reilly, 1969).

Individuals with ASD have been shown to have low variability in their behaviour (Baron-Cohen, 1992; Miller & Neuringer, 2000). This is likely to be related to the prevalence of RRBs in this population and can have implications for an individual's performance across a variety of repertoires, including play, language, problem-solving and reinforcement maximisation (Mullins & Rincover, 1985; Neuringer, 2002). This in turn has important implications for successful performance in social, vocational and academic settings.

Through careful experimentation and using larger sample sizes than previous research, the current work aimed to establish the extent of the differences in variability in responding between individuals with ASD and their neurotypical peers using a variety of measurement techniques, including the U-value statistic. As much of our understanding of the mechanisms underlying operant variability and the consequential control that can alter its frequency has its foundation in research with non-human animals, the current thesis aimed to extend the applications of this previous work by applying such procedures to a clinical population. The participant

groups involved in this work were children and adolescents who were neurotypical and children and adolescents who had a diagnosis of ASD. The neurotypical participants served as comparison groups for the ASD participants.

Another important aim of the current work was to investigate the relationship between behavioural variability and severity and type of restricted repetitive behaviours in children and adolescents with ASD. It also aimed to identify some of the variables associated with low variability in this population, including age, severity of ASD symptoms, language ability and play skills.

The final aim of the current thesis was to apply the findings from previous experimental studies to a clinically relevant behaviour, variability in play actions with toys, by identifying optimal conditions to induce a high level of variability. This is the first study of its type that has actual play actions as the target behaviour. Previous studies targeted variability in the product of the behaviour, e.g., the structure built with blocks (Napolitano, Smith, Zarcone & McAdam, 2010) or in selection of toys or activities (Camilleri & Hanley, 2005). It aimed to increase variability in play in a meaningful way among the clinical population in order to identify procedures with real benefits that can be incorporated into intervention programmes. The extension of procedures from basic to applied research is important here. The focus of such procedures in the past has been in research with non-human animals and, while more applied studies are beginning to emerge (e.g., Susa & Schlinger, 2012), there has been little examination of the precise conditions that promote behavioural variability with human participants.

The importance of play in the development of early social behaviour has been well documented (Sutton-Smith, 1975; Wenner, 2009). Children with ASD have

been shown to have low variability in their play behaviours and they often demonstrate repetitive actions with toys (Stahmer & Schreibman, 1992). This is likely to result in serious impairments in their ability to interact with their peers through play, which in turn may negatively impact the acquisition of important social skills. The development of techniques to increase variability in play skills is considered an important contribution to interventions for children with ASD.

## 6. 2. Summary of Thesis Findings

In Chapter 2, the first empirical chapter of this thesis, a large sample (n = 130) was utilised to investigate the difference in variability between children and adolescents with ASD and those of typical development. Robust conclusions were made that supported the previous findings of Baron-Cohen (1992) and Miller and Neuringer (2000). Participants with ASD demonstrated significantly less variability in responding than their neurotypical peers. The procedure incorporated the Penny-Hiding Game which was appropriate for this purpose as it is commonly observed in naturalistic interactions. Interestingly, nine participants with ASD selected the same hand across the 18 trials of the game. No neurotypical participants demonstrated this pattern of responding. A large effect size was observed for the between-groups difference in this study.

This study is unique as it is the first time that low variability has been investigated in the context of RRBs. To achieve this aim, variability in individuals with ASD was related to the presence and type of RRB using the Repetitive Behaviour Scales – Revised (Bodfish et al. (2000). As expected, there was a significant negative correlation between variability and total RBS-R score, indicating

that lower variability was strongly related to high rates of RRBs in this population (n = 65). When relationships between variability and RRB subtypes were investigated, an interesting finding was highlighted. Significant negative correlations were found between variability and Stereotyped Behaviour and between variability and Self-Injurious Behaviour; these belong to the class of RRBs labelled Repetitive Sensory and Motor Behaviours (RSMBs). These are described as lower order RRBs and are generally related to greater impairment in functioning (Lam, Bodfish & Piven, 2008).

The subtypes of the RBS-R that correspond to higher order RRBs (Ritualistic/Sameness Behaviour, Compulsive Behaviour and Restricted Interests) were not found to be significantly correlated with variability, indicating that individuals who present with these types of RRBs have greater ability to vary their behaviour than those presenting with lower order RRBs. This study contributes to the existing knowledge base on operant variability by revealing important information about the complex relationship between low variability and RRBs in this population.

There have been some conflicting reports in the experimental and applied literature around the effect of age on variability (Bancroft, 2011; Neuringer, 2002). In the current study, the data revealed no significant effect of age on variability.

The difference in variability between neurotypical and children with ASD was investigated in Chapter 3 also, this time using a carefully controlled method of data collection. A sophisticated and automated data collection system was developed for the purpose of assessing variability in responding using the U-value statistic. This statistic was an important component of the study as it provides a clear indicator of the uncertainty of the distribution of responses, with values closer to 0 being

stochastic and those closer to 1 being random. This also allowed for careful comparison across individuals and groups. A computer program in the form of a game was designed for this purpose. Furthermore, this study utilised an age equivalence matched control group as well as a chronological age matched control group in an attempt to control for the effect of cognitive functioning on variability in responding following the outcomes reported in Chapter 2. The Peabody PPVT-4 was used to determine age equivalence scores of the participants with ASD in the interest of efficiency and given that it provides a a standardised measure of receptive language (formally referred to as “mental age” and often used as a measure of non-verbal IQ).

Statistical analysis revealed that there was a significant difference in variability between the participants with ASD and the participants matched for age equivalence. There was a difference between participants with ASD and participants matched for chronological age but it did not reach statistical significance. Importantly, in each case, the neurotypical participants demonstrated higher variability than those with ASD.

In this study, there was a significant negative correlation between age and variability in responding, with younger participants demonstrating greater variability. No such effect was observed in the previous chapter. It is not clear whether it was the precise measurement utilised in Chapter 3 that resulted in this finding or if it was related to other, unmeasured participant characteristics. There was no significant relationship between variability and performance on the PPVT-4, indicating that language ability (as provided by a standardised measure of receptive language) was not related to variability in this sample.

Chapter 4 utilised a specifically designed and more sophisticated computer program in the form of a computerised game that was designed for the purpose of the study. The development of this resource ensured greater motivation for participants as they interacted with a more complex and appealing interface, using a series of animated characters. This program also allowed for a more complex presentation of a series of lag schedules. The program consisted of four different settings, each corresponding to a different sequence of presentation of the lag schedules, thus controlling for sequence effects.

In Chapter 4, lag schedules of reinforcement were incorporated within the game in order to investigate their effects on variability in responding. As before, a U-value statistic was calculated as part of the program as a measure of variability, resulting in quick and easy comparison of variability across different conditions, participants and groups. Furthermore, a control condition, where reinforcement was independent of variability, was incorporated into the program along with the following lag conditions: lag 1, lag 2, lag 4, lag 6, lag 8. The control condition was important as it allowed the experimenter to demonstrate that the lag schedules were responsible for the observed variability in responding.

Significant differences in variability were found between the control condition and the lag 6 condition and the control condition and the lag 8 condition, indicating that the greatest increases in variability were observed in the presence of the higher lag values. The difference between the control condition and the lag 4 condition was significant at the $p < .05$ level but when the Bonferroni correction was applied it did not reach significance. Nonetheless, it is useful to note that the lag 4 condition also resulted in a high level of variability.

As part of Study 3, outcomes of the computer-based test of variability were examined in light of the participants' scores on the ADOS-G. There was some indication that there was an inverse relationship between ADOS-G scores and U-values in this sample, i.e., participants with greater severity of ASD symptoms demonstrated lower variability. However, there was variation within the sample and no conclusive remarks can be made at this time. Interestingly, the participant who had the highest score on the ADOS-G (greatest autism severity) demonstrated the lowest variability across the intervention and he achieved a U-value of 0 in the lag 2 condition. This means that, even though there was a lag 2 schedule of reinforcement in place, this participant repeated the same sequence of key presses across all 96 trials in that condition. It is possible that variability and responsiveness to lag schedules is affected by autism severity only at extreme ends of the spectrum but this warrants further investigation. The current findings serve as an interesting starting point for such investigation.

Furthermore, Study 3 yielded outcomes of six participants demonstrating high and stable variability across the conditions (four from the NT group and two from the ASD group). This represents approximately one quarter of the sample. It is not clear at this time whether the design of the computerised game contributed to this or not. It may be the case that the intrinsic motivation embedded within the game led the participants to try as many options as possible from the outset in order to progress through it. As mentioned in Chapter 4, Abreu-Rodrigues, Lattal, Dos Santos & Matos, (2005) reported very similar findings for one of their four pigeons when using lag schedules. As this was not the primary focus of the research the authors did not offer a robust explanation for this phenomenon. However, if human patterns are found to reflect those observed in the animal research then perhaps we can learn

much more from those studies in terms of patterns of results and susceptibility to procedures that may produce meaningful change in the applied field.

Chapter 4 once again demonstrated the differences in variability between the participants with ASD and the neurotypical participants, with the U-value statistic producing a robust demonstration of the phenomenon. This study developed a sophisticated manner of determining such a difference and the U-value was an important part of this demonstration as it allowed for quick and easy comparison across individuals, groups and conditions. While the overall U-values were lower among the participants with ASD, the performances shown in response to the lag schedules was very similar. Inspection of the graphs of the mean U-values for each condition of each of the groups reveals the same shape across the conditions with the highest variability being observed in the lag 6 condition and slightly lower variability observed in the lag 8 condition. This indicates that there may be a ceiling effect of variability when schedules are increased past a lag 6 when using human participants.

Chapter 5 was comprised of two applied experiments where the findings of the previous empirical chapters were used to develop a procedure for increasing variability in a socially significant behaviour, play actions with toys, in an applied setting. This chapter was significant as the research on increasing variability in play behaviours has been confined to block building (Goetz & Baer, 1973; Napolitano, Smith, Zarcone, Doodkin & McAdam, 2010) and has been limited to lag 1 schedules of reinforcement (Camilleri & Hanley, 2005; Napolitano et al. 2010).

Lee, Sturmey and Fields (2007) highlight the specific shortcomings of a lag 1 reinforcement schedule whereby “fixed variability requirements” (p. 439) may result in stereotyped alternations between certain patterns of responding. Simply

alternating between two responses repeatedly on a lag 1 schedule meets the variability requirement for reinforcement. We note that Schwartz (1982) refers to this pattern of responding as a higher-order stereotypy. In this sense, use of a lag 1 schedule with a clinical population, particularly those with autism diagnoses, may inadvertently reinforce forms of higher-order steroptyped patterns of responding within the particular topographical target behaviour. The two studies reported within Chapter 5 provided an important step in their aims to address the limitations reported with lag 1 schedules in previous research (Lee, McComas & Jawor, 2002; Lee & Sturmey, 2006). Effectively, increased lag reinforcement schedules were explored to determine their effect on repetition of response patterns. The current study incorporated increased complexity in both the target behaviour under study and the procedures used.

The two experiments in Chapter 5 utilised an array of age-appropriate toys that were found to be appealing to neurotypical children. Thus, it was thought that they would be approximately representative of the toys that the participants would encounter in inclusive play or school settings. Prior to the intervention in Experiment 1, two neurotypical children were observed separately playing with the toys and an action list was devised following these observations. The neurotypical children showed high variability in their actions with the toys and reported afterwards that they enjoyed playing with them. Interestingly, in Experiment 1, three of the five participants showed little interest in the toys during the baseline phase and demonstrated very few play actions with them. One participant refused to engage with them in any way. Following the lag reinforcement procedure, all five participants demonstrated an increase in the number of different actions they demonstrated with the toys relative to baseline. This was considered an important

and positive outcome of the procedure. Some four of the five participants demonstrated a clear upward trend across the training phase, indicating that the lag schedule continued to increase variability in play actions. While four of the five participants emitted fewer actions when the schedule was removed than when it was in place, a maintenance effect was still observed for all participants when compared to baseline.

Important secondary findings of this study were an increase in the number of toys sampled even though this was not directly targeted. Furthermore, for the two participants that demonstrated a high frequency of stereotypy at baseline, this decreased significantly during and after the intervention, indicating that the increase in variability in actions resulted in a concurrent decrease in stereotypy. This is a very significant finding as it provides support for the suggestions of Lee et al. (2007) and Napolitano et al. (2010) that methods for increasing variability in responding may have important implications for the treatment of stereotypy in the clinical population. The presence of RRBs in individuals with ASD is likely to be due to restricted repertoires. By prompting and shaping new behaviours, we can ensure that more naturally occurring reinforcers are encountered in the environment. Lee et al. (2007) report that the greater the number of variations of responses that contact contingencies of reinforcement, the more variable the individual's responding will become. In this way, the mechanisms that are responsible for response allocation, for example, the matching law, may also be responsible for variability in behaviour patterns.

Pre- and post- baseline probes were conducted with a set of generalisation toys to determine if the increase in variability in actions would generalise to new stimuli. For four of the five participants, an increase was observed, indicating good

generalisation as a function of the lag procedure. Napolitano et al. (2010) reported poor generalisation of variability in block-building following their lag 1 reinforcement procedure. Furthermore, the training stimuli and the generalisation stimuli were quite similar (wooden blocks vs. plastic blocks). In Chapter 5, Experiment 1, an entirely different toy set was used to measure generalisation. Thus, the positive findings described above serve as a strong indication of generalisation for the procedure employed in this thesis.

Interestingly and in light of Lee et al.'s (2007) discussion of the limitation of lag schedules in producing fixed variability requirements resulting in stereotyped alterations between response forms, one participant (Participant 4) in Experiment 1 demonstrated this particular response pattern. The participant fulfilled the requirements of the lag schedule but he did so while demonstrating a repetitive chain of actions. Schwartz (1982) described such responding as higher-order stereotypy and this type of responding, under lag reinforcement schedules, may be more obvious with an autism population given the prevalence of repetitive behaviour observed in this population by definition of the diagnosis. In integrated settings this pattern of play behaviour would be considered atypical and it highlights an important consideration when using lag schedules to treat restricted behavioural repertoires in applied settings.

In Experiment 1 (Chapter 5) a procedure was developed that allowed for an efficient and effective way to increase variability in actions with toys for five participants with ASD. The five participants presented with a range of autism severity, indicating that the procedure was effective regardless of the participant characteristics at baseline. While more sessions may have been required to achieve high variability and generalisation for the participants who were lower functioning,

the method was robust across participants. The efficacy of this procedure in the applied setting was another positive finding of the current study.

Experiment 2 (Chapter 5) isolated one participant from Experiment 1 in order to examine the effects of a more rigorous single-case design and demonstrate the effects of further exposures to the lag reinforcement schedules. A reversal design was used to demonstrate that the lag schedule, and not just exposure to the toys, was responsible for the increase in variability. In this experiment, the lag value was increased incrementally from lag 1 to lag 2 to lag 3 across the three phases. Different toys were used to control for learning and satiation effects of previous phases in Experiment 1. The participant demonstrated increasing variability as the lag values were increased. He also demonstrated increased variability with the generalisation toys, even though the lag schedule was never used with these. The positive outcomes shown with an incremental increase in the lag schedule on play behaviour were promising and they provide sound evidence for the efficacy of using such schedules in applied settings. Although this may be considered a preliminary investigation, the two experiments reported in Chapter 5 provide strong support for the procedure which was developed for this purpose.

The analysis of high frequency actions conducted in Experiment 2 revealed significant findings. The participant demonstrated several actions at a very high rate in the lag 1 phase. By the time he reached the criterion in the lag 3 phase, there were very few actions that occurred at a high frequency. It represented a more equal spread across the play actions. This indicated that increasing the lag requirement served as a remediation strategy for repetition of actions at a high rate.

Chapter 5 represents a novel area of investigation with more complex strategies than previous work in applying lag schedules to restricted behavioural repertoires in autism. It shows how the strategies that were developed resulted in meaningful behaviour change within an important repertoire for the five participants with ASD.

## 6. 3. Theoretical Implications of the Current Findings

The current thesis has several important theoretical implications. Firstly, it showed that there is a significant difference in variability between children with ASD and their neurotypical peers, with the children with ASD consistently demonstrating low variability in responding. This was related for the first time to the presence of restricted repetitive behaviours, specifically the lower order category of repetitive sensory and motor behaviours (RSMBs).

It was revealed in Chapter 3 that the higher lag values (lag 4, lag 6 and lag 8) resulted in the greatest variability in responding. While this had been demonstrated in published research with college students (Balsam, Deich, Ohyama & Stokes, 1998) and with non-human animals (Abreu-Rodrigues et al., 2005; Page & Neuringer, 1985), the current thesis replicated these results with neurotypical children and children with ASD. Importantly, this reveals that even lower functioning children who demonstrate extremely low variability and high rates of RRBs show sensitivity to changing lag schedules.

Another important theoretical implication is that lag schedules can be used to increase variability in complex behaviours involving an array of stimuli. In

Experiment 1 (Chapter 5) there was a range of 16 toys available. The rationale for this was that it would more closely emulate a typical naturalistic setting where children play. This further supports the notion that lag schedules can be moved "out of the laboratory" and have efficacy in less controlled settings with more complex behaviours. From a practical point of view, investigating conditions that ameliorate response variability in "real word" settings should aid in the development and refinement of novel procedures that decrease or eliminate RRBs among a clinical population.

The generalisation effects observed in Experiments 1 and 2 (Chapter 5) are another important theoretical implication of the current thesis. Prior to this, the capacity of lag schedules for producing variability that would generalise was relatively unknown. Goetz (1982) reported that generalisation effects would be likely when the transfer task was somewhat similar to the training task. In Chapter 5, this was shown to be the case. While a novel set of toys was used, many similar actions could be demonstrated with them such as opening and closing doors, driving vehicles, placing figurines inside objects etc. Lee et al. (2007) describe response covariation as instances where systematic changes in one response result in systematic changes in other response that have not been directly targeted. They suggest that variability in one response class may influence the amount of variability in another response class. The current findings indicate that this may be possible in children with autism, particularly when different values of lag schedules are used.

The observed decrease in stereotypy following the implementation of lag schedules for increasing variability in play actions raises an important theoretical question. It is unclear whether the reduction in stereotypy was due to an increase in variability that affected the low variability actions, i.e., stereotypy. Alternatively, it

is possible that the combination of response interruption and redirection (RIRD) and reinforcement of varied actions acted as a DRI for stereotypy. Emitting varied actions with toys is indeed incompatible with most motor stereotypies; however, further investigation is warranted in order to identify the conditions under which the greatest and most lasting decreases in stereotypy can be observed. Nonetheless, the decrease in stereotypy for two participants who were engaging at a very high rate at baseline is a significant outcome of the current thesis with important implications for how researchers and clinicians understand stereotypy and other RRBs in the context of behavioural variability.

Derby et al. (1997) implemented functional communication training (FCT) with four children with developmental delays and they noted that when they trained a mand that was functionally equivalent to an aberrant behaviour, the aberrant behaviour decreased and manding increased. However, a variety of appropriate behaviours, not directly targeted also increased. The authors concluded that as the aberrant behaviours decreased and manding increased, other forms of appropriate behaviour contacted reinforcement and overall, there was more reinforcement available due to the increases in appropriate behaviour. The decrease in stereotypy observed in the current study may be a result of a variety of alternative responses encountering reinforcement.

### 6. 4. Implications for Autism Intervention

The findings of the current thesis have important implications for intervention for individuals with ASD. The observed relationship between low variability and high rates of RRBs indicates that strategies that increase variability in responding are likely to be successful in decreasing RRBs in this population. As discussed above,

the exact mechanisms for bringing about this change require further study but the current outcomes indicate that using lag schedules of reinforcement to increase variability in play actions with toys can result in a marked decrease in stereotypy during play.

The most significant implication of the current thesis for autism intervention is the development and testing of an efficient and effective procedure for increasing variability in play skills. When variability increased, the participants demonstrated many novel actions with the toys. Thus, they were trying new things in order to fulfil the lag schedule. Furthermore, they engaged with the toys in more meaningful way and they also sampled more of the available toys. Lee et al. (2007) report that an increase in appropriate behaviour leads to a "greater overall density of reinforcement" (p. 446). Individuals with ASD may require more reinforcement than their peers to shape their behaviour and also to encounter reinforcement for a wider range of responses. Increasing variability in a socially significant repertoire such as play actions with toys may result in the emergence of further new response classes due to the increased density of reinforcement in the environment. This has important implications for the ability of these participants to integrate into inclusive settings.

As demonstrated in Chapter 4, large increases in variability can be achieved with relatively few training sessions that can be of short duration. The efficacy of using higher lag values with the target behaviour of variability in play actions with toys was unknown prior to the current study. The procedure outlined in Chapter 4 could easily be incorporated into a daily programme for children with ASD. It remains to be seen if this type of procedure will yield optimal outcomes in early intervention settings where toys involve many concrete movements, or whether the

procedures could also be used with older adolescent children and more complex play materials.

The current findings add to the existing research on using lag schedules with participants with ASD. They suggest that when a high level of variability is required, higher lag values should be used. The practicality of implementing higher lag values in applied settings may be questioned due to the level of monitoring of responses or sequences of responses. However, Lee et al. (2007) describe the advantages of the monitoring process involved in lag schedules:

"The degree of recency is specified in the lag value. Thus, only responses that fall within the lag value are monitored. The greater the lag value, the greater the set of responses that are ineligible for reinforcement" (p. 442).

When using lag schedules to increase variability in verbal behaviour, for example in tact or mand frames, it should be possible to implement higher lag values and monitor the specified dimension (e.g., topography) from the sample of previous behaviour with relative ease. The data from Chapter 4 indicate that a lag 6 schedule will yield a high level of variability. However, increasing the lag value past this may result in an extinction of variability so this should be avoided. Another method for utilising higher lag values is to incorporate automated programs into ASD programme design. However, the extent to which this will generalise to increased variability in socially significant behaviours such as play and verbal behaviour is not yet clear.

The effective use of prompting in Experiment 1 (Chapter 5) indicates that prompting can be incorporated into procedures involving lag schedules. In this case, the use of prompts allowed for quick increases in variability which is desirable in

applied settings. Least to most prompting should always be used, with a focus on fading the prompts as quickly as possible, thereby shifting the control from the prompts to the schedule itself.

## 6. 5. Implications for Future Research

The findings of the current thesis highlight several important areas for future research that will further our understanding of operant variability and the prevalence of low variability in individuals diagnosed with ASD. The relationship between variability and a variety of factors, including autism severity, cognitive functioning and adaptive functioning warrants investigation with a large sample of individuals with ASD across a wide age range. Furthermore, the prevalence and impact of low variability on adults with ASD is an under-researched area which may reveal important findings.

The participant characteristics that influence responsiveness to interventions involving lag schedules of reinforcement also warrant further investigation. For example, in Chapter 4 of the current study there were four participants in the neurotypical group and two participants in the ASD group who had high and stable variability across the six lag conditions. Future research may reveal the underlying characteristics that result in such outcomes.

Further investigation involving the use of lag schedules could consider the use of a variable lag schedule as suggested by Lee, Sturmey and Fields (2007). In Experiment 1 (Chapter 4), Participant 4 demonstrates what these and other authors (Schwartz, 1982) described as higher order stereotypy. In contrast to higher order

RRBs described by Turner (1999) and Lam et al. (2008), Lee et al. (2007) use the term to refer to chains of responding that fulfil the lag requirement but do so in a stereotyped way. For example, when Participant 4 was exposed to a lag 3 schedule of reinforcement, he emitted a sequence of four different actions that he repeated in a chain, thus obtaining reinforcement for each action while actually engaging in quite repetitive responding. This was the only participant who was observed to respond in this way but a remediation strategy for this may be to vary the lag value within conditions as opposed to across conditions.

Susa and Schlinger (2012) achieved an increase in variability in tacting using a lag 3 schedule of reinforcement. Unlike the experiments outlined in Chapter 5, they implemented the lag 3 schedule directly following the baseline phase. This reveals an area of research that may have a large applied significance. Investigations that determine the most efficient way of using lag schedules may assist in refining such procedures for effective use in applied settings. Such research should compare increases in variability when the lag value is increased systematically, as in the current thesis, and when higher lag values are implemented directly following baseline, as in the Susa and Schlinger (2012) study. A multiple baseline design across participants or repertoires of behaviour would likely be useful for such an analysis.

The utility of lag schedules for increasing variability in responding has been demonstrated in the repertoire of language and play. Additional research is warranted to expand the research base to incorporate such procedures across different repertoires such as problem-solving (in both academic and real-life examples) and exploration of new environments. Future research could also expand on the current thesis and investigate the use of lag schedules for increasing variability in more

complex play and social behaviours e.g., interactive play. The development of strategies for implementation of the procedure and measuring outcomes would assist in their becoming more widely utilised. The development of software, such as apps for smartphones may lead to important advancements in the way in which such procedures are incorporated into intervention programmes for individuals with ASD.

The two experiments in Chapter 4 incorporated generalisation probes and they revealed promising outcomes. However, as noted above, the transfer task (novel toys) could be considered similar to the training task. Further investigation is warranted to determine the extent of generalisation effects as a function of lag schedules. Topographically dissimilar tasks would be useful for this purpose. Furthermore, generalisation across settings should also be investigated. Finally, it would be useful to determine the optimum lag schedule value and amount of exposure to the schedule that yields the greatest generalisation of variability in responding.

## 6. 6. Limitations of the Current Thesis

**6.6.1. Sampling and generality of findings.** As is commonly the case when using human participants, there are limitations to the current thesis regarding sampling of the population. As all of the participants required parental consent in order to participate, it is likely that the type of parents who readily consent to their child's involvement in research studies share some common characteristics that may influence certain characteristics in their children. However, as informed parental consent was paramount to each child's participation in this work, this potential bias was unavoidable.

In the two experiments of Chapter 4, all of the participants attended an ABA school and had done so for more than one year at the time of implementation of the experiments. Because of this, they were familiar with the use of token economies, which was an advantage to the experimenter. However, if there had been participants who had never encountered a token economy, direct reinforcement could have been used or, alternatively, a procedure for introducing a token economy (including conditioning of tokens) could have been conducted prior to the intervention. The extent to which the success of the intervention with the students with ASD who attended an ABA school can be generalised to students with ASD who have not received ABA intervention is unclear but it is thought that the outcomes were related to participant characteristics such as autism severity rather than type of intervention.

**6.6.2. Standardised Measures Used.** Chapter 2 utilised the RBS-R, an empirically-validated standardised measure of restricted repetitive behaviours. However, it is an indirect method of assessment and may result in some degree of subjectivity on the part of the individuals completing it. In Chapter 3, the control group matched for age equivalence was sampled according to the age equivalence scores obtained by the participants with ASD on the PPVT-4. A standardised IQ test such as the WISC-IV may have been more suitable as previous work and the outcomes of Chapter 2 suggested that cognitive functioning was likely to be related to variability in responding. However, time constraints imposed by the schools in which the research was conducted meant that it was not practical to administer the WISC-IV. The PPVT-4 was a favourable compromise as it is quick to administer and has been used in previous published studies as a measure of non-verbal IQ (e.g., Reed, Broomfield, McHugh, McCausland & Leader, 2009).

**6.6.3. Maintenance and Follow-Up.** The current thesis demonstrated positive outcomes of the procedures used. However, due to the practical constraints of time and access to participants, it was not possible to test for maintenance effects of these procedures. Because of this, conclusions cannot be drawn relating to the long-term benefits of such procedures. It is unclear whether ongoing training would be required in order to achieve maintenance and generalisation effects over time. Further research is required to determine the extent of maintenance from the type of procedure described in this thesis and if further training across multiple repertoires would be required.

## 6. 7. Conclusion

The current thesis adds to the existing research base on operant variability by extending previous work from the Experimental Analysis of Behaviour. Many important findings resulted from the current work that have implications for our understanding of low variability in individuals with a disgnosis of ASD and how to remediate this. A significant relationship was found between low variability and high overall rates of RRBs and, in particular, the class of RRBs referred to as RSMBs. This is a new finding that may change how we view the instances of RRBs in this population. Furthermore, when lag schedules were implemented to increase variability in toy play, the concurrent reduction in stereotypy was an exciting finding.

The current investigation into the use of lag schedules with individuals with ASD and neurotypical individuals also yielded important findings. Firstly, they were effective in increasing variability in both populations but it must be considered that the neurotypical participants demonstrated a high level of variability, even in the

control condition where reinforcement was independent of variability. It was discovered that when a significant increase in variability is desired, higher lag values (lag 4 or greater) should be used where possible.

The use of lag schedules with individuals with ASD in applied settings is supported by the outcomes of the current thesis. It also shows that lag schedules can be used to increase variability in more complex play behaviours. While previous work in this area was restricted to increasing variability in block building tasks and selection of activities, the current work increased variability in complex play actions with toys, indicating the efficacy of lag schedules for increasing variability across a variety of target behaviours.

Furthermore, significant increases in variability can be achieved in a short time and the procedures are efficient to implement. This is of benefit to practitioners designing programmes for individuals with ASD. The outcomes reported in this thesis provide support for the operant nature of variability and demonstrate that variability can be influenced by consequences. The contingent reinforcement of variable responding can be used to increase variability across a range of repertoires. This is a novel area of research but it is one of huge applied significance. Procedures that increase behavioural variability in individuals with ASD may have a large role to play in the development of new interventions that can remediate the low variability commonly observed in this population and facilitate the ability to communicate and socailise more effectively with their peers. Ulitmately, it is hoped that such procedures will be incorporated into intervention programmes and this will lead to improvements in the ability of individuals with autism to integrate successfully into inclusive social, academic, vocational and recreational settings.

## References


Abreu-Rodrigues, J., Lattal, K.A., Dos Santos, C.V., & Matos, R.A. (2005). Variation, repetition and choice. *Journal of the Experimental Analysis of Behaviour. 83,* 147-168. doi: 10.1901/Fjeab.2005.33-03

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders (4th Edition)*: Washington D.C.

Antonitis, J. J. (1951). Response variability in the white rat during conditioning, extinction, and reconditioning. *Journal of Experimental Psychology, 42,* 273-281. doi: 10.1037/h0060407

Atlas, J.A. (1990). Play in assessment and intervention in the childhood psychoses. *Child Psychiatry and Human Development, 21,* 119-133. doi:10.1007/BF00706120

Balsam, P. D., Deich, J. D., Ohyama, T., & Stokes, P. D. (1998). Origins of new behavior. In W. O'Donohue (Ed.), *Learning and behavior therapy* (pp. 403-420). Boston: Allyn & Bacon.

Bancroft, S. (2011). Behavioral Variability in Selection of Play Items: Assessment, Treatment, and Generalization. *Proquest Dissertations & Theses.* http://proquest.umi.com/pqdlink?did=2564634301&Fmt=7&clientId=79356&RQT=309&VName=PQD

Baron-Cohen, S. (1992). Out of sight or out of mind? Another look at deception in autism. *Journal of Child Psychology & Psychiatry, 33,* 1141-1155. doi: 10.1111/j.1469-7610.1992.tb00934.x

Becker, R.F. (1981). Creativity and problem solving: An overview. Retrieved from http://digital.library.wisc.edu/1793/6084

Betz, A. M., Higbee, T. S., Kelley, K. N., Sellers, T. P., & Pollard, J. S. (2011). Increasing response variability of mand frames with script training and extinction procedures. *Journal of Applied Behavior Analysis. 44.* 357- 362. doi: 10.1901/jaba.2011.44-357

Blough, D. S. (1966). The reinforcement of least frequent interresponse times. *Journal of the Experimental Analysis of Behavior, 9,* 581-591. doi:10.1901/jeab.1966.9-581

Bodfish, J.W., Symons, F.J., & Lewis, M.H. (1999). The Repetitive Behavior Scales: A test manual. Morgantown, NC: *Western Carolina Research Reports.*

Bodfish, J.W., Symons, F.J., Parker, D.E. & Lewis, M.H., (2000). Varieties of

repetitive behavior in autism: Comparisons to mental retardation. *Journal of Autism and Developmental Disorders, 30.* 237-243. doi: 10.1023/A:1005596502855

Bowman, L. G., Piazza, C. C., Fisher, W. W., Hagopian, L. P., & Kogan, J. S. (1997). Assessment of preference for varied versus constant reinforcers. *Journal of Applied Behavior Analysis, 30,* 451-458. doi: 10.1901/jaba.1999.32-437

Boucher, J. (1977). Alternation and sequencing behaviors and response to novelty in autistic children. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 18,* 66-72. doi: 10.1111/j.1469-7610.1977.tb00417.x

Bransford, J. D., Arbitman-Smith, R., Stein, B. S., & Vye, N. J. (1985). Analysis-improving thinking and learning skills: An analysis of three approaches. In S. F. Chipman, J.W. Segal, & R. Glaser (Eds.), Thinking and learning skills: Vol. 1 (pp. 133-206). Hillsdale, NJ: Erlbaum.

Cammilleri, A. P. & Hanley, G. P. (2005). Increasing varied selections of classroom activities. *Journal of Applied Behavior Analysis, 38,* 111-116. doi: 10.1901/jaba.2005.34-04

Carr E.G., & Kologinsky E. (1983). Acquisition of sign language by autistic children. II: Spontaneity and generalization effects. *Journal of Applied Behavior Analysis, 16,* 297–314. doi: 10.1901/jaba.1983.16-297

Charman, T., & Baron-Cohen, S. (1997). Prompted pretend play in autism. *Journal of Autism and Developmental Disorders, 27,* 325-332.

Cherot, C., Jones, A., & Neuringer, A. (1996). Reinforced variability decreases with approach to reinforcers. *Journal of Experimental Psychology: Animal Behavior Processes, 22,* 497-508. doi: 10.1037/0097-7403.22.4.497

Cooper, J.O., Heron, T.E., & Heward, W.L. (2007). *Applied Behavior Analysis (2nd Ed).* New Jersey: Pearson Education Inc.

Cunningham, A.B., & Schreibman, L. (2008). Stereotypy in autism: The importance of function. *Research in Autism Spectrum Disorders, 2,* 469-479. doi: 10.1016/j.rasd.2007.09.006.

Denney, J., & Neuringer, A. (1998). Behavioral variability is controlled by discriminative stimuli. *Animal Learning & Behavior, 26,* 154-162. doi: 10.3758/BF03199208

DiLavore, P., Lord, C., & Rutter, M. (1995). Pre-Linguistic Autism Diagnostic

Observation Schedule (PL-ADOS). *Journal of Autism and Developmental Disorders, 25,* 355–379.

Driver, P. M., & Humphries, D. A. 1988. *Protean Behaviour: The Biology of Unpredictability*. Oxford: Clarendon Press.

Duker, P.C., & Van Lent, C. (1991). Inducing variability in communicative gestures used by severely retarded individuals. *Journal of Applied Behavior Analysis, 24,* 379-386. doi: 10.1901/jaba.1991.24-379

Dunn, L. M., & Dunn, D. M**.** (2007). *Peabody Picture Vocabulary Test, Fourth Edition.* Minneapolis, MN: NCS Pearson.

Dunn, L. M., Dunn, L. M., Whetton, C., & Pintilie, D. (1982). British Picture Vocabulary Scale. Windsor, Berks: NFER Nelson.

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. doi: 10.1901/jeab.1969.12-73

Eisenberger, R., & Selbst, M. (1994). Does reward increase or decrease creativity? *Journal of Personality & Social Psychology, 66,* 1116-1127. doi: 10.1037/0022-3514.66.6.1116

Epstein, R. (1985b). Extinction-induced resurgence: Preliminary investigations and possible applications. *The Psychological Record, 35,* 143-153.

Epstein, R. (1996). *Cognition, creativity and behavior: Selected essays*. Westport, CT: Praeger.

Gabriels, R.L., Agnew, J.A., Miller, L.J., Gralla, J., Pan, Z., Goldson, E., Ledbetter, J.L., Dinkins, J.P., & Hooks, E. (2008). Is there a relationship between restricted, repetitive, stereotyped behaviors and interests and abnormal sensory response in children with autism spectrum disorders. *Research in Autism Spectrum Disorders, 2,* 660-670. doi: 10.1016/j.rasd.2008.02.002

Gabriels, R.L., Cuccaro, M.L., Hill, D.E., Ivers, B.J., & Goldson, E. (2005). Repetitive behaviors in autism: Relationships with associated clinical features. *Research in Developmental Disabilities, 26,* 169-181. doi: 10.1016/j.ridd.2004.05.003

Goetz, E. M. (1982). A review of functional analyses of preschool children's creative behaviors. *Education and Treatment of Children, 5,* 157-177.

Goetz, E. M., & Baer, D. M. (1973). Social control of form diversity and emergence

of new forms in children's blockbuilding. *Journal of Applied Behavior Analysis, 6,* 209-217. doi: 10.1901/jaba.1973.6-209

Gorlitz, D. (1987). Exploration and attribution in a developmental context. In D. Gorlitz & J. F. Wohlwill (Eds.), *Curiosity, imagination and play: On the development of spontaneous, cognitive and motivational processes.* Hillsdale, NJ: Erlbaum.

Grossman, J.B., Carter, A., & Volkmar, F.R. (1997). Social behaviour in autism. *Annals of the Ney York Academy of Sciences, 807,* 440-454. doi: 10.1111/j.1749-6632.1997.tb51938.x

Grunow, E.M., & Neuringer, A. (2002). Learning to vary and varying to learn. *Psychonomic Bulletin & Review, 9,* 250-258. doi: 10.3758/BF03196279

Harding, J.W., Wacker, D.P., Berg, W.K., Rick, G., & Lee, J.F. (2004). Promoting response variability and stimulus generalization in martial arts training. *Journal of Applied Behaviour Analysis, 37,* 185-195. doi: 10.1901/jaba.2004.37-185

Heldt, J. & Schlinger, H.D. (2012). Increased variability in tacting under a lag 3 schedule of reinforcement. *The Analysis of Verbal Behaviour, 28,* 131-136.

Holman, J., Goetz, E. M., & Baer, D. M. (1977). The training of creativity as an operant and an examination of its generalization characteristics. In Etzel, B. LeBland, J., & Baer, D. (Eds.), *New developments in behavior research: Theory, method and application* (pp. 441-471). Hillsdale, NJ: Erlbaum.

Hopkinson, J. & Neuringer, A.(2003). Modifying behavioral variability in moderately depressed students. *Behavioral Modification: 27,* 254-261. doi: 10.1177/0145445503251605

Jarrold, C., Boucher, J., & Smith, P. (1993). Symbolic play in autism: A review. *Journal of Autism and Developmental Disorders, 23,* 281-387.

Jones, R.S.P., (1990).The social effects of stereotyped behaviour. *Journal of Mental Deficiency Research, 34,* 261-268. doi: 10.1111%2Fj.1365-2788.1990.tb01537.x

Koegel, R.L., O'Dell, M.C. and Koegel, L.K. 1987: A natural language teaching paradigm for nonverbal autistic children. *Journal of Autism and Developmental Disorders* **17**, 187–200.

Lalli, J.S., Zanolli, K., & Wohn, T. (1994). Using extinction to promote response

variability in toy play. *Journal of Applied Behavior Analysis, 27,* 735-736. doi: 10.1901/jaba.1994.27-735

Lam, K.S.L. (2004). The Repetitive Behavior Scale-Revised: Independent validation and the effects of subject variables. Unpublished Doctoral Dissertation: Ohio State University.

Lam, K.S.L., & Aman, M.G. (2007). The Repetitive Behaviour Scale – revised: Independent Validation in Individuals with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 33,* 855-866. doi: 10.1007/s10803-006-0213-z

Lam, K. S. L., Bodfish, J.W., & Piven, J. (2008). Evidence for three subtypes of repetitive behaviour in autism that differ in familiality and association with other symptoms. *Journal of Child Psychology and Psychiatry, 49,* 1193-1200. doi:10.1111/j.1469-7610.2008.01944.x

Landis, J.R.; & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics* **33** (1): 159–174. doi:10.2307/2529310

Lee, R., McComas, J.J., & Jawor, J. (2002). The effects of differential and lag reinforcement schedules on varied verbal responding by individuals with autism. *Journal of Applied Behavior Analysis, 35,* 391-402. doi: 10.1901/jaba.2002.35-391

Lee, R., & Sturmey, P. (2006). The effects of lag schedules and preferred materials on variable responding in children with autism. *Journal of Autism and Developmental Disorders, 36,* 421-428.

Lee, R., Sturmey, P., & Fields, L. (2007). Schedule-induced and operant mechanisms that influence response variability: A review and implications for future investigations. *The Psychological Record, 57,* 429-455

Leslie, A. M. (1987). Pretence and representation: the origins of "theory of mind". *Psychological Review, 94,* 412-426. doi:10.1037%2F0033-295X.94.4.412

Lewis, V., & Boucher, J. (1995). Generativity in the play of young people with autism. *Journal of Autism and Developmental Disorders, 25,* 105-121. doi:10.1007/BF02178499

Lewis , M.H., Tanimura, Y. , Lee, L.W. , & Bodfish, J.W. (2007). Animal models of restricted repetitive behavior in autism. *Behavioural Brain Research, 176*(1), 66-74 . doi:10.1016/j.bbr.2006.08.023

Lewis, M.H., & Bodfish, J.W. (1998). Repetitive behavior disorders in autism. *Mental Retardation and Developmental Disabilities Research Reviews, 4*, 80-89. doi:10.1002/(SICI)1098-2779(1998)4:2<80::AID-MRDD4>3.0.CO;2-0

Libby, S., Powell, S., Messer, D., & Jordan, R. (1998). Spontaneous play in children with autism: A reappraisal. *Journal of Autism and Developmental Disorders, 28*, 487-497.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism Diagnostic Observation Schedule: A standardized observation of communicativeand social behavior. *Journal of Autism and Developmental Disorders, 19*, 185–212.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview–Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659–685.

Lord, C., Risi, S., Cook, E.H. Jr., Leventhal, B.L., DiLavore, P.C., Pickles, A., Rutter, M. (2000). The autism diagnostic observation schedule – generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30*, 205-223.

Lovaas, O.I., Schreibman, L., Koegel, R., & Rehm, R. (1971). Selective responding by autistic children to multiple sensory input. *Journal of Abnormal Psychology, 77*, 211-222. doi:10.1037/h0031015

Machado, A. (1989). Operant conditioning of behavioral variability using a percentile reinforcement schedule. *Journal of the Experimental Analysis of Behavior, 52*, 155-166. doi: 10.1901/jeab.1989.52-155

Maltzman, I. (1960). On the training of originality. *Psychological Review, 67*, 229-242. doi: 10.1037/h0046364

Manoel, E. de J. & Connolly, K.J. (1995). Variability and the development of skilled actions. *International Journal of Psychophysiology. 19*, 129-147. doi: 10.1016/0167-8760(94)00078-S

Mayer, R.E., & Greeno, J.G. (1972). Structural differences between learning outcomes produced by different instructional methods. *Journal of Educational Psychology, 63*, 165-173. doi:10.1037%2Fh0032654

Mayfield, K.H., & Chase, P.N. (2002). The effects of cumulative practice on mathematical problem-solving. *Journal of Applied Behavior Analysis, 35*,

105-123. doi: 10.1901/jaba.2002.35-105

Miller, G.A., & Frick, F.C. (1949). Statistical behavioristics and sequences of responses. *Psychological Review, 56,* 311-324. doi: 10.1037/h0060413

Miller, N., & Neuringer, A. (2000). Reinforcing variability in adolescents with autism. *Journal of Applied Behavior Analysis, 33,* 151-165. doi: 10.1901/jaba.2000.33-151

Mullins, M., & Rincover, A. (1985). Comparing autistic and normal children along the dimensions of reinforcement maximization, stimulus sampling, and responsiveness to extinction. *Journal of Experimental Child Psychology, 40,* 350-374. doi: 10.1016/0022-0965(85)90095-5

Mundy, P., Sigman, M., Ungerer, J., & Sherman, T (1986). Defining the social deficits of autism: The contribution of nonverbal communication measures. *Journal of Child Psychology and Psychiatry, 27,* 657-669. doi: 10.1111%2Fj.1469-7610.1986.tb00190.x

Napolitano, D.A., Smith, T., Zarcone, J.R., Goodkin, K., & McAdam, D.B. (2010). Increasing response diversity in children with autism. *Journal of Applied Behaviour Analysis, 43,* 265-271. doi: 10.1901/jaba.2010.43-265

Neuringer, A. (2002). Operant variability: Evidence, functions and theory. *Psychonomic Bulletin & Review, 9,* 672-705. doi: 10.3758/BF03196324

Neuringer, A. (2004). Reinforced variability in animals and people. *American Psychologist, 59,* 891-906. doi: 10.1037/0003-066X.59.9.891

Neuringer, A., Deiss, C., & Olsen, G. (2000). Reinforced variability and operant learning. *Journal of Experimental Psychology: Animal Behavior Processes. 26,* 98-111. doi: 10.1037/0097-7403.26.1.98

Neuringer, A., & Huntley, R.W. (1992).Reinforced variability in rats: Effects of gender, age and contingency. *Physiology & Behaviour, 51,* 145-149. doi: 10.1016%2F0031-9384%2892%2990216-O

Neuringer, A., & Jensen, G. (2012). The predictably unpredictable operant. *Comparative Cognition and Behaviour Reviews, 7,* 55-84. doi: 10.3819/ccbr.2012.70004

Neuringer, A., Kornell, N., & Olufs, M. (2001). Stability and variability in extinction. *Journal of Experimental Psychology: Animal Behavior Processes, 27,* 79-94. doi: 10.1037//0097-7403.27.79

Nevin, J. A. (1967). Effects of reinforcement scheduling on simultaneous

discrimination performance. *Journal of the Experimental Analysis of Behavior*, *10*, 251-260. doi: 10.1901/jeab.1967.10-251

Oswald D.P., Ollendick T.H. (1989). Role taking and social competence in autism and mental retardation. *Journal of Autism and Developmental Disorders. 19*, 119-127.

Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*, 429-452. doi: 10.1037/0097-7403.11.3.429

Paterson, R. C., & Arco, L. (2007). Using video modeling for generalizing toy play in children with autism. *Behaviour Modification*, *31*, 660-681. doi: 10.1177%2F0145445507301651

Pellegrini, A. D. (1995). *School Recess and Playground Behaviour*. Albany: State University of New York Ed.

Pellegrini, A.D., Dupuis D., and Smith P.K., (2007). Play in evolution and development. *Developmental Review 27*, 261-276. doi: 10.1016/j.dr.2006.09.001

Pierce, K., & Courchesne, E. (2001). Evidence for a cerebellar role in reduced exploration and stereotyped behaviour in autism. *Biological Psychiatry*, *49*, 655-664. doi:10.1016/S0006-3223(00)01008-8

Powell, S., & Jordan, R. (21/4/2004). Teaching Individuals with Autism Flexible Thinking and Autonomous Behaviour. Retrieved from http://www.autismo-br.com.br/home/Artig083.htm.

Prior, M., & MacMillan, M. B. (1973). Maintenance of sameness in children with Kanner's syndrome. *Journal of Autism and Childhood Schizophrenia*, *3*, 154-167. doi:10.1007%2FBF01537990

Pryor, K.W., Haag, R., & O'Reilly, J. (1969). The creative porpoise: Training for novel behavior. *Journal of the Experimental Analysis of Behavior*, *12*, 653-661. doi: 10.1901/jeab.1969.12-653

Rapp, J.T., & Vollmer, T.R. (2005). Stereotypy I: A review of behavioral assessment and treatment. *Research in Developmental Disabilities*, *26*, 527-547. doi: 10.1016/j.ridd.2004.11.005

Reed, P., Broomfield, L., McHugh, L., McCausland, A., & Leader, G. (2009).

Extinction of over-selected stimuli causes emergence of under-selected cues in higher-functioning children with autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 39,* 290–298

Rincover, A., Cook, R., Peoples, A., & Packard, D. (1979). Sensory extinction and sensory reinforcement principles for programming multiple adaptive behavior change. *Journal of Applied Behavior Analysis, 12,* 221–233. doi: 10.1016/j.biopsych.2007.02.024

Rincover, A. & Koegel, R.L. (1975). Setting generality and stimulus control in autistic children. *Journal of Applied Behavior Analysis, 8,* 235-246. doi: 10.1901/jaba.1975.8-235

Rogers, V.R., & Ghezzi, P.M. (2010). Extinction-induced variability in young children with autism. *36th Annual Convention of Applied Behavior Analysis International*. San Antonio, Texas.

Ross, C., & Neuringer, A. (2002). Reinforcement of variations and repetitions along three independent response dimensions. *Behavioral Processes, 57,* 199-209. doi: 10.1016/S0376-6357(02)00014-1

Roeyers, H., & van Berkalaer-Onnes, I. (1994). Play in autistic children. *Communication and Cognition, 27,* 349-359.

Ruscio, A.M., & Amabile, A.M. (1999). Effects of instructional style on problem-solving creativity, *Creativity Research Journal, 12,* 251-266. doi: 10.1207%2Fs15326934crj1204_3

Rydell, P., & Prizant, B. M. (1995). Assessment and intervention strategies for children who use echolalia. In Quill, K. (Ed.), *Teaching children with autism: Strategies to enhance communication and socialization.* Albany, NY: Delmar.

Saldana, R. L., & Neuringer, A. (1998). Is instrumental variability abnormally high in children exhibiting ADHD and aggressive behavior? *Behavioral Brain Research, 94,* 51-59. doi: 10.1016/S0166-4328(97)00169-1

Schoenfeld, W.N., Harris, A.H., & Farmer, J. (1966). Conditioning response variability. *Psychological Reports, 19,* 551-557. doi: 10.2466%2Fpr0.1966.19.2.551

Schwatrz, B. (1982b). Reinforcement induced behavioural stereotypy: How not to teach people to discover rules. *Journal of the Experimental Analysis of Behaviour, 111,* 23-59. doi: 10.1037%2F%2F0096-3445.111.1.23

Shahan, T.A., & Chase, P.N. (2002). Novelty, stimulus control and operant

variability. *The Behaviour Analyst, 25,* 175-190.

Siegler, R.S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology, 28,* 225-273. doi: 10.1006/cogp.1995.1006

Skinner, B.F. (1938) *The Behavior of Organisms: An Experimental Analysis*. New York: Appleton Century.

Spinka, M., Newberry, R.C., & Beckoff, M. (2001). Mammalian play: Training for the unexpected. *The Quarterly Review of Biology. 76,* 141-168. doi:10.1086/393866

Stahmer, A. C. (1995): Teaching symbolic play to children with autism using pivotal response training. *Journal of Autism and Developmental Disorders* **25**, 123–41. doi: 10.1007%2FBF02178500

Stahmer, A., Ingersoll, B. & Carter, C. (2003). Behavioral approaches to promoting play. *Autism – International Journal of Research and Practice, 7,* 401-414.

Stahmer, A.C., & Schreibman, L. (1992). Teaching children with autism appropriate play using a self-management treatment package. *Journal of Applied Behavior Analysis, 25,* 447-459. doi: 10.1901/jaba.1992.25-447

Stokes, P.D. (1995). Learned variability. *Animal Learning and Behavior, 23,* 164-176. doi: 10.3758/BF03199931

Stokes, P. D., Mechner, F., & Balsam, P. D. (1999). Effects of different acquisition procedures on response variability. *Animal Learning & Behavior, 27,* 28-41. doi: 10.3758/BF03199429

Susa, C.L., & Schlinger, H.D. (2012). Using a lag schedule to increase variability of verbal responding in an individual with autism. *The Analysis of Verbal Behaviour. 28,* 125–130.

Sutton-Smith, B. (1975). The useless made useful: Play as variability training. *School Review, 83,* 197-214. doi:10.1086/443187

Turner, M.A. (1999) Annotation: Repetitive behavior in autism: A review of psychological research. *Journal of Child Psychology and Psychiatry, 40,* 839-849. doi: 10.1111/1469-7610.00502

Ungerer, J.A., & Sigman, M. (1981). Symbolic play and language comprehension in autistic children. *Journal of the American Academy of Child Psychiatry, 20,* 318-337. doi:10.1016/S0002-7138(09)60992-4

Wenner, M. (2009). The serious need for play, *Scientific American Mind, 25,* 22-29

**APPENDIX A**
**Action List - Chapter 5: Study 4, Experiment 1**

| **TOY** | **ACTION** |
|---|---|
| Dinosaur | Open mouth |
| | Move arm |
| | Move leg |
| | Bang head off item |
| | Walk |
| | Hop |
| | Place items in mouth |
| | Place item on dino's back |
| | Take item of back |
| | Stamp on item |
| Cave | Remove top |
| | Open door |
| | Close door |
| | Stack rooms |
| | Place items inside cave |
| | Take item out of cave |
| Swing | Move swing |
| | Place item in swing |
| | Take item out of swing |
| | Take off swing |
| | Put on swing (hang) |
| Caveman Car | Place item in car |
| | Take item out of car |
| | Roll in straight line |
| | Roll in curved line |
| | Fly through air |
| | Crash car into object |
| Egg | Open egg |
| | Close egg |
| | Place item in egg |
| | Take item out of egg |
| | Place egg in new location |
| Cave man | Walk along |
| | Hop along |
| | Climb |
| Cave woman | Walk along |
| | Hop along |
| | Climb |

| | |
|---|---|
| Baby | Walk along<br>Hop along<br>Climb |
| Baby Dino | Walk along<br>Hop along<br>Climb |
| Police Station | Open door<br>Close door<br>Ring telephone (press button)<br>Place items inside<br>Take items out |
| Police car | Place item inside car<br>Take item out of car<br>Roll in straight line<br>Roll in curved line<br>Fly through air<br>Crash car into object |
| Police man | Walk along<br>Hop along<br>Climb |
| Police woman | Walk along<br>Hop along<br>Climb |
| Police dog | Walk along<br>Hop along<br>Climb |
| Traffic light | Move location |
| Stop sign | Move location |