THE EFFECTS OF INTERRESPONSE INTERVALS ON BEHAVIORAL VARIABILITY IN HUMANS

THESIS

Presented to the Graduate Council of the University of North Texas in Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

By

Mark P. Reilly, B. S.
Denton, Texas
December, 1993

The present experiment studied the relationship between interresponse intervals and behavioral variability. Subjects emitted sequences of 4 keypresses on two keys on a variability schedule that delivered points when the current 4-response sequence differed from the previous 5 sequences. Three experimental conditions were studied; no interresponse interval, 4-s interresponse interval and 8-s interresponse interval. Interresponse intervals followed each of the first three responses in each sequence. Two groups were used to study initial training histories. Group 1 was first exposed to the no-interresponse interval condition. Group 2 was first exposed to the 4-s interresponse interval condition. Subjects were then exposed to the different interresponse interval conditions.

There was little change in variability across conditions. However, the variability observed in the subjects first exposed to the 4-s interresponse interval was greater than the variability observed in subjects first exposed to no-interresponse interval. There was higher-order response patterning in both groups, but it was more pronounced in the no-interresponse interval group.
I would like to acknowledge my major professor, Sigrid Glenn for her valuable guidance through this project. Also, I am greatly indebted to my committee members, Joel Greenspoon, Cloyd Hyten and Janet Ellis for their support. Finally, I would like to thank Patricia Benetz for her patience and support.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>v</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF ILLUSTRATIONS</td>
<td>vi</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. METHOD</td>
<td>6</td>
</tr>
<tr>
<td>Subjects</td>
<td></td>
</tr>
<tr>
<td>Apparatus</td>
<td></td>
</tr>
<tr>
<td>Procedure</td>
<td></td>
</tr>
<tr>
<td>III. RESULTS</td>
<td>11</td>
</tr>
<tr>
<td>Session 1 analysis</td>
<td></td>
</tr>
<tr>
<td>Effects of conditions</td>
<td></td>
</tr>
<tr>
<td>IV. DISCUSSION</td>
<td>46</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>51</td>
</tr>
</tbody>
</table>
LIST OF TABLES

TABLE 1.  Number of sessions in each condition ....................... 10
TABLE 2.  Means and standard deviations of percent correct and U-values . 13
TABLE 3.  Subjects’ percent correct and U-values for training session ...... 30
LIST OF ILLUSTRATIONS

FIGURE 1. Mean percent correct as a function of IRI .......... 12
FIGURE 2. Mean U-value as a function of IRI ............ 15
FIGURE 3. Percent correct and U-values for JXP ............ 17
FIGURE 4. Percent correct and U-values for CXP ............ 18
FIGURE 5. Percent correct and U-values for MRG .......... 19
FIGURE 6. Percent correct and U-values for AXL .......... 20
FIGURE 7. Percent correct and U-values for MXD .......... 21
FIGURE 8. Percent correct and U-values for SDG .......... 22
FIGURE 9. Event records of training session for Group 2 subjects .... 25
FIGURE 10. Event records of training session for Group 1 subjects .... 26
FIGURE 11. Correlograms of training session for Group 2 subjects .... 27
FIGURE 12. Correlograms of training session for Group 1 subjects .... 28
FIGURE 13. Event records for JXP across conditions ......... 31
FIGURE 14. Event records for CXP across conditions ......... 32
FIGURE 15. Event records for MRG across conditions ......... 33
FIGURE 16. Event records for AXL across conditions ......... 34
FIGURE 17. Event records for MXD across conditions ......... 35
FIGURE 18. Event records for SDG across conditions ......... 36
FIGURE 19. Correlograms for JXP across conditions ......... 40
FIGURE 20. Correlograms for CXP across conditions ................. 41
FIGURE 21. Correlograms for MRG across conditions .................. 42
FIGURE 22. Correlograms for AXL across conditions .................. 43
FIGURE 23. Correlograms for MXD across conditions .................. 44
FIGURE 24. Correlograms for SDG across conditions .................. 45
CHAPTER I

INTRODUCTION

Behavioral variability, as an outcome resulting from certain types of selection contingencies, has received considerable attention lately within behavior analysis (Machado, 1989, 1992; Morgan & Neuringer, 1990; Morris, 1987, 1989, 1990; Neuringer, 1991; Page & Neuringer, 1985). The general finding is that behavioral variability can be generated reliably in nonhumans when certain contingencies of reinforcement are used. Machado (1989, 1992) has discussed these contingencies as frequency-dependent selection contingencies in which the least frequent or least probable sequences are the most likely to be reinforced. Behavioral variability is treated as an outcome of these selection contingencies.

In a typical procedure used to study response variability in pigeons, two keys, a left (L) and a right (R), are transilluminated the same color. Pigeons are trained to peck the lit key, regardless of its position. Following training, both keys are lit, and a schedule is in effect that delivers food for keypecking the transilluminated keys when the current 4-keypeck sequence differs from the previous N sequences, where N is usually constant and is called the "lag" value (e.g., Page & Neuringer, 1985). When N = 5, the current sequence has to be different from the previous 5 sequences to produce food. For example, food would not be delivered following an LLRR sequence on a Lag 5 schedule if the previous sequences were RLLR, LLRR, LLLL,
RRLL, and LRRR. Food would be delivered following an RRRR sequence in this instance.

This procedure consists of discrete-trials because between responses in each sequence, the keylights are off for a designated interresponse interval (IRI). Discrete-trial procedures have been successful in generating variable response sequences (Cohen, Neuringer, & Rhodes, 1990; Machado, 1989, 1992; Morgan & Neuringer, 1990; Morris, 1987, 1989, 1990; Neuringer, 1991; Page & Neuringer, 1985). There is evidence that discrete-trial procedures, as opposed to procedures that have no IRI between responses of each sequence, increase the likelihood that sequence variability will be produced (Morris, 1987, 1989, 1990; Neuringer, 1991; Page & Neuringer, 1985; Schwartz, 1980, 1982).

Schwartz (1982) used a procedure without IRI’s between the first three responses of each sequence in an experiment designed to study behavioral variability in pigeons. Although reinforcement was contingent on variability, the pigeons developed stereotypic response patterns and obtained the minimal amount of the total reinforcers that could have been earned during each session. Schwartz’s conclusion was that reinforcement is, by theoretical definition and by empirical demonstration, a mechanism that increases response stereotypy. He concluded that reinforcement could not select variable patterns of responding. This conclusion has been shown incorrect as the literature on behavioral variability provides many examples of variability that were determined by the selecting contingencies established by the experimenter (Machado, 1989, 1992; Morgan & Neuringer, 1990; Morris, 1987, 1989, 1990;
Neuringer, 1991; Page & Neuringer, 1985). The failure of Schwartz's pigeons to respond variably was due to the specific features of his procedures (Catania, 1987; Morris, 1987, 1989; Page & Neuringer, 1985). First, Schwartz's experiments used a free-operant procedure with no IRI's; second, trials consisted of sequence lengths of 8 with the added constraint that no more than 4 responses on each key could occur and meet the contingency. This reduced the total possible sequences from 256 to 70.

Although either of these procedural features may explain the negative results obtained by Schwartz, the free-operant procedure has been proposed as the more likely explanation (Morris, 1987, 1989, 1990; Page & Neuringer, 1985).

Morris (1987, 1989) attempted to determine the differences between free-operant and discrete-trial procedures by comparing their effects on behavioral variability in pigeons on a variability schedule with a lag of 2. The important difference between the two procedures was the presence or absence of the IRI. The results of the 1987 study showed that pigeons received less than one third of the available reinforcers in the free-operant (No-IRI) procedure. Pigeons in the discrete-response procedure, on the other hand, obtained three-fourths of the available reinforcers. Similar results were obtained in a later replication with greater lag values (Morris, 1989). Morris concluded that the discrete-trial procedure, which separated each response with a brief blackout interval, was more effective in producing response variability than the free-operant procedure that lacked an IRI. We are still left with the unanswered question of what is it about the nature of the IRI that increases response variability.
Morris (1987, 1989, 1990) argued that Pavlovian processes may be interfering with the control of behavioral variability in the free-operant procedure. His argument was that the keylights elicit pecks and thus inhibit the development of behavioral variability under the free-operant procedure. Because the IRI in the discrete-trial procedure is essentially a blackout period, the conditioned stimulus (lights) cannot elicit keypecking, and therefore, keypecking may come under the control of operant reinforcement (Morris, 1990). An operant explanation is also possible.

One difference between discrete-trial and free-operant procedures is the response rates they generate. It may be that response rate is a key factor in generating behavioral variability. An inverse relationship may exist between response rate and variability. Simply stated, as response rates increase, behavioral variability decreases; conversely, when response rates decrease, behavioral variability increases. Although this relationship has not received direct study, it has been observed and reported in several previous studies (Antonitis, 1951; Cohen, Neuringer, & Rhodes, 1990; Eckerman & Lanson, 1969; Herrnstein, 1961; Neuringer, 1991).

Neuringer (1991) studied the effects of different IRI values on sequence variability. Pigeons were exposed to the following order of IRI values, 0.5, 2.0, 0.1, 0.3, 1.0, 0.0, and 4.0 s across conditions in which keypecking on two keys was reinforced on a variability schedule with a lag of 5. Each IRI value was in effect for five sessions. The results showed sequence variability (as measured by percent correct and U-value) increased as IRI values increased. Neuringer concluded that sequence variability increased because longer IRIs interfered with the ability to
remember previous response sequences. Another explanation is that response rates were manipulated indirectly by IRI manipulations. Perhaps response rate affects variable performance by determining whether responses exert control over the following responses in a 4-response sequence. Maybe a chain is created when responses follow one another rapidly. Each response would seem more likely to acquire discriminative function for the next response. This chain would be less likely to form when greater intervals of time separate responses. This may account for the variability observed by Neuringer (1991).

The present experiment attempted to replicate Neuringer’s 1991 experiment. Humans served as subjects, instead of rats, and different initial training conditions were provided for 2 groups of subjects. Group 1 subjects were trained initially with the free-operant procedure (No-IRI condition), and Group 2 subjects were trained initially with the discrete-response procedure in which 4-s IRIs separated the first 3 responses in each sequence (4-s IRI condition). Next, subjects were exposed to 1 or 2 other IRI conditions (all subjects were exposed to an 8-s IRI condition). The main question asked by the present study was: What is the relationship between IRI and behavioral variability in humans and how does it compare with nonhuman findings? Moreover, is there a difference in behavioral variability between humans trained initially on discrete-response (IRI condition) and on free-operant (No-IRI condition) procedures? And are there any order effects caused by exposures to the two conditions?
CHAPTER II

METHOD

Subjects

Six college students were selected as subjects based on their availability to participate for the length of the experiment. Applicants responded to an advertisement in the college newspaper recruiting students interested in participating for money in a learning experiment.

Apparatus

The experiment took place in a small room that contained a table, a chair, and an 80386 personal (IBM compatible) computer. Subjects were seated in front of a computer keyboard and video monitor. Two 8 cm x 8 cm squares were centered on the video monitor. They were placed 4 cm apart. The left square was blue, and the right square was green. The two shift keys on the keyboard served as operand and were colored blue and green to correspond to the two stimuli appearing on the video monitor. The shift keys were the only keys functional on the keyboard. Experimental manipulations, stimulus presentations, and data analysis were done using Pascal programming language. Subjects sometimes were viewed from an adjacent room through a one-way mirror.
Procedure

A single-subject design were used to compare the effects of 3 IRI conditions on behavioral variability. Six undergraduate subjects were pseudorandomly assigned to one of two groups. Group 1 subjects received their initial training in the No-IRI condition, and Group 2 subjects received their initial training in the IRI condition with IRI equal to 4 s. The rest of the procedures followed Neuringer’s 1991 Experiment 1 and are described below.

Subjects pressed sequences of four left and right shift keys on a computer keyboard. When a key was pressed, the corresponding square on the monitor turned grey and moved 1 cm to the right. The square returned to its original color and position upon release of the key when the No-IRI condition was in effect. In the IRI condition, the entire screen blacked out for 4 or 8 s (depending on which condition was in effect) when the key was released. In both conditions, the fourth response in each sequence was followed either by point delivery or an intertrial interval (ITI), depending on whether the current sequence met the reinforcement contingency. If the current sequence differed from the previous 5 sequences, 2 points were displayed on the computer monitor for 3 s accompanied by a tone that sounded at 440-Hz and increased in frequency. If the current sequence occurred anywhere in the previous 5 sequences, the fourth response in that sequence was followed by an ITI of 3 s in which the video monitor was darkened, and two tones of 220 and 150 Hz occurred sequentially. Any responses occurring during this ITI reset the 3-s interval.
There was a training session for all subjects that was identical to the above procedure except that it had a lag value of 1 (the current sequence had to be different from the immediately preceding sequence). Two points were delivered each time the current sequence differed from the immediately preceding sequence. Group 1 subjects began in the No-IRI condition (free-operant). The No-IRI condition had no blackout following the first 3 responses of each 4-response sequence. Group 2 subjects began in the IRI condition (discrete-trial), in which each of the first 3 keypresses in a sequence of 4 keypresses was followed by a 4-s IRI during which the screen was blacked-out. Responses during the IRI reset the 4-s interval. Before the training session, subjects were read the following instructions:

Your job is to earn as many points as possible. Points will be exchangeable for money at the end of each session.

To earn points, press the SHIFT keys when the colored boxes are on the screen. Use only your preferred finger to press the keys. Each point is worth 0.02 cent.

Stay in the room until the computer indicates that the session is over.

These written instructions remained in the room throughout the remainder of the experiment. Following this training session, the variability contingency was increased to Lag 5. Two points were delivered every time the current sequence differed from the previous 5 sequences.

The present study investigated the effects of different initial training conditions and of 4 and 8-s IRI conditions versus no IRI conditions on the degree of variability obtained. To examine the effects of the initial training conditions, two groups of subjects were given different initial training conditions. The procedures were the
same except for the order of conditions, which differed so as to counterbalance and elucidate any order effects caused by the different histories of exposure. Each condition was in effect for four consecutive sessions unless otherwise indicated.

Following the four sessions in one condition, subjects were exposed to a different IRI condition for four more sessions. Subjects then were returned to either the original condition, or a different condition for four sessions. Table 1 shows the order and number of sessions in each condition that each subject received. IRI and No-IRI conditions were manipulated such that each subject received an exposure to at least 2 different conditions. The present experiment was initially proposed as a reversal design, but because of small changes observed in variability across the No-IRI and the 4-s IRI conditions, some subjects received an exposure to an 8-s IRI condition rather than a return to baseline. Others received an 8-s IRI as the only IRI condition.

Sessions were conducted daily, 7 days a week. All sessions ended when 100 points were delivered except for sessions under the 8-s IRI condition. These sessions were shortened to 50 points to improve the likelihood that subjects remained in the experiment. Each point was worth $0.02 in each condition except for the 8-s IRI condition, in which points were worth $0.04. This kept the amount earned constant. Subjects were paid in full after each session. To make it likely that subjects completed the experiment proper, a lottery system was set up so that all subjects who completed the experiment had an equal chance of winning $30 (an exception was subject MRG who started after everyone else had completed the experiment). At the end of the experiment, each subject was asked to describe how they earned money
and what they did during the different conditions during a debriefing session with the experimenter.

Table 1

Number of sessions in each condition for each subject

<table>
<thead>
<tr>
<th>Subject</th>
<th>Condition</th>
<th>Sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>MXD</td>
<td>No IRI</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>8-s IRI</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>No IRI</td>
<td>4</td>
</tr>
<tr>
<td>AXL</td>
<td>No IRI</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>4-s IRI</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>8-s IRI</td>
<td>4</td>
</tr>
<tr>
<td>SDG</td>
<td>No IRI</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>8-s IRI</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>No IRI</td>
<td>3</td>
</tr>
<tr>
<td>JXP</td>
<td>4-s IRI</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>No IRI</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>8-s IRI</td>
<td>4</td>
</tr>
<tr>
<td>CXP</td>
<td>4-s IRI</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>No IRI</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>8-s IRI</td>
<td>3</td>
</tr>
<tr>
<td>MRG</td>
<td>4-s IRI</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>No IRI</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>8-s IRI</td>
<td>3</td>
</tr>
</tbody>
</table>
CHAPTER III

RESULTS

The traditional measures of response variability were taken and included percentage of "correct" sequences (or percentage of sequences that met the contingency) and U-values. Percentage of sequences that meet the contingency is an indication of how well the subject's behavior conforms to the variability contingency. The higher the percentage, the greater the control exerted by the contingency. This value was calculated by dividing the total number of sequences that were followed by points by the total number of sequences emitted during each session. Figure 1 shows percent correct as a function of IRI. Each symbol represents the mean of the last 3 sessions of each subject's first exposure to the 3 different conditions. As shown in Figure 1, five subjects earned fewer of the possible points as IRI increased. The decrease in percent correct, although systematic, was slight. Table 2 shows the means and standard deviations of percent correct and U-values of the data plotted in Figures 1 and 2. Three of the 6 subjects, MRG, MXD and CXP showed at least an 7% decrease in percent correct from the No-IRI to the 8-s IRI condition. Subjects MXD and MRG had the biggest differences between percent correct at the No-IRI and the 8-s IRI conditions with 17.07 and 17.83, respectively. One subject, SDG, earned more potential points as IRI was increased to 8 s (SDG's percentages were 80.7 in the No-IRI condition and 84 in the 8-s IRI Figure 1 condition). There were no clear
Figure 1. Mean percent correct as a function of interresponse interval (IRI) for last 3 sessions in each condition. Closed symbols are subjects that had training in the No-IRI condition (Group 1). Open symbols are subjects that had initial training in the 4-s IRI condition (Group 2).
Table 2

Means and standard deviations of percent correct and U-values for the last three sessions in each condition for all subjects

<table>
<thead>
<tr>
<th>Subject</th>
<th>Condition</th>
<th>No IRI</th>
<th></th>
<th>4-s IRI</th>
<th></th>
<th>8-s IRI</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>%</td>
<td>SD</td>
<td>%</td>
<td>SD</td>
<td>%</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>M</td>
<td></td>
<td>M</td>
<td></td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>JXP</td>
<td></td>
<td>94.90</td>
<td>.51</td>
<td>90.96</td>
<td>2.82</td>
<td>88.76</td>
<td>.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95.64</td>
<td>.18</td>
<td>95.54</td>
<td>.20</td>
<td>94.86</td>
<td>.75</td>
</tr>
<tr>
<td>CXP</td>
<td></td>
<td>94.03</td>
<td>1.32</td>
<td>88.93</td>
<td>4.90</td>
<td>86.77</td>
<td>3.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95.69</td>
<td>.10</td>
<td>94.26</td>
<td>1.11</td>
<td>95.03</td>
<td>.17</td>
</tr>
<tr>
<td>MRG</td>
<td></td>
<td>87.56</td>
<td>3.88</td>
<td>72.63</td>
<td>1.09</td>
<td>69.73</td>
<td>5.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>74.72</td>
<td>8.26</td>
<td>91.60</td>
<td>1.17</td>
<td>83.40</td>
<td>4.50</td>
</tr>
<tr>
<td>AXL</td>
<td></td>
<td>92.63</td>
<td>2.88</td>
<td>92.03</td>
<td>1.96</td>
<td>87.30</td>
<td>3.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td>71.80</td>
<td>0.00</td>
<td>73.60</td>
<td>2.36</td>
<td>71.66</td>
<td>.22</td>
</tr>
<tr>
<td>MXD</td>
<td></td>
<td>91.80</td>
<td>2.98</td>
<td>N/A</td>
<td>N/A</td>
<td>74.73</td>
<td>19.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>72.43</td>
<td>01.63</td>
<td>N/A</td>
<td>N/A</td>
<td>74.53</td>
<td>3.80</td>
</tr>
<tr>
<td>SDG</td>
<td></td>
<td>80.77</td>
<td>6.36</td>
<td>N/A</td>
<td>N/A</td>
<td>84.00</td>
<td>5.16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>73.20</td>
<td>4.06</td>
<td>N/A</td>
<td>N/A</td>
<td>87.00</td>
<td>4.21</td>
</tr>
</tbody>
</table>
differences in percent correct across conditions between the subjects in the 2 groups.

In addition to percent correct, U-value is a common measurement of response variability (Page & Neuringer, 1985). U-value is defined as

\[
\sum_{i=1}^{16} \left[ p_i \times \frac{\log(p_i)/\log(2)}{\log(n)/\log(2)} \right] \times 100
\]

where \( p \) equals the obtained probability of each sequence represented by \( i \), and \( n \) equals the number of possible sequences, which is 16 in the present case. The U-value is a molar measure of the degree of variability among all the possible sequences (Miller & Frick, 1949, as reported by Page & Neuringer, 1985). It is based on the obtained probabilities of each of the 16 possible sequences. If \( U = 0 \), 1 sequence (e.g., LLLR) was repeated throughout the session. If \( U = 100 \), obtained probabilities for each of the 16 sequences are equal and prediction of a specific sequence may be impossible because each sequence is equally likely.

Figure 2 shows U-values as a function of IRI. Each symbol in Figure 2 represents the mean of the last three sessions of the first exposure to each IRI condition. As can be seen, four out of six subjects' U-values remained relatively constant across all three IRI conditions. However, the values differed depending on
Figure 2. Mean U-value as a function of interresponse interval (IRI) for last 3 sessions in each condition. Closed symbols are subjects that had initial training in the No-IRI condition (Group 1). Open symbols are subjects that had initial training in the 4-s IRI condition (Group 2).
the initial condition. The two subjects in Group 2 stayed consistently above a 92 U-value, and the two subjects in Group 1 stayed consistently around 70. Two subjects, MRG and SDG, however, showed higher U-values in the IRI conditions. Subject SDG's U-values increased from 73.2 at No IRI to 87.0 at 8-s IRI. MRG's U-values began at 91.6 in the 4-s IRI condition and decreased to 74.7 in the No-IRI condition and then increased to 83.4 at the 8-s IRI condition. Unlike percent correct in Figure 1, Figure 2 shows some differences between the two groups. Specifically, Group 2 subjects' (open symbols) U-values were substantially higher than the U-values of Group 1 (closed symbols), indicating higher overall variability among the 16 possible sequences for Group 2. As seen in the percent correct and the U-values in Figures 1 and 2 then, there is relatively little change in variability as a function of IRI conditions.

It is possible, also, to examine the dynamics of the subjects' behavior as it occurs over time across sessions and experimental manipulations. Figures 3, 4, and 5 show percent correct (upper graphs) and U-values (lower graphs) across sessions for subjects in Group 2 (4-s IRI initial exposure). For all three subjects, percent correct appears slightly higher in the No-IRI conditions, and lower in the 8-s IRI conditions although these differences are not great. Figure 5 (subject MRG) shows the biggest change in percent correct. Some of the data (CXP and MRG) appear unstable indicating possible trends. Too few sessions were conducted to draw any firm conclusions. U-values for JXP and CXP are all 90 or greater and are virtually unchanged across conditions. Subject MRG (Figure 5) again shows the greatest
Figure 3. Percent correct across sessions (top) and U-values across sessions (bottom) for JXP.
Figure 4. Percent correct across sessions (top) and U-values across sessions (bottom) for CXP.
Figure 5. Percent correct across sessions (top) and U-values across sessions (bottom) for MRG.
Figure 6. Percent correct across sessions (top) and U-values across sessions (bottom) for AXL.
Figure 7. Percent correct across sessions (top) and U-values across sessions (bottom) for MXD.
Figure 8. Percent correct across sessions (top) and U-values across sessions (bottom) for SDG.
change across conditions. MRG’s U-value decreased by 20% during the No-IRI condition from the constant U-value around 95% during the initial 4-s IRI condition.

Figures 6, 7, and 8 show percent correct and U-values for subjects in Group 1 (No-IRI initial exposure). All three subjects displayed increasing trends in percent correct during their first exposure to the No-IRI condition. The percent correct for each subject (except SDG) was highest in the No-IRI condition and lowest in the 8-s IRI condition. For SDG, percent correct was consistently higher in the 8-s IRI condition than it was in the first exposure to the No-IRI condition. The second exposure to the No-IRI condition produced higher percent correct than the 8-s IRI condition. The U-values in Figures 6 (AXL) and 7 (MXD) all are in the 70 range and display little change across conditions. For subject SDG in Figure 8, the U-values showed an increasing trend regardless of the condition. This subject’s percent correct (except for the second exposure to the No-IRI condition) and U-values most closely approximate the nonhuman percent correct and U-values in Neuringer’s 1991 experiment.

Training session analysis

To aid understanding of the differences in U-values as a function of initial exposure to the contingencies, we shall consider performance during the training session. Recall that the training session had a Lag 1 variability contingency that delivered points if the current 4-response sequence was different from the immediately preceding sequence. One way to look at what the subject did during a session is to identify the sequences by numbering them arbitrarily and plot which sequences were
emitted consecutively throughout the session. Since there were 16 possible 4-response sequences distributed over two keys, each sequence was given an identifying number (ID). These numbers were then plotted as they occurred across consecutive sequences in something like an event record. This way of presenting the data shows two forms of variability; variability amongst the 16 possible sequences (the bandwidth of the scatterplots) and variability in terms of periodicity (the randomness or scatter of the data points across consecutive sequences). The former is frequently reported in the variability literature as a statistical measure such as the U-value which was presented above. The latter form of variability addresses serial dependency and is infrequently reported in the variability literature (see however, Machado, 1992).

Figures 9 and 10 present training session data plotted in this way. Figure 9 is the training session event record for subjects in Group 2 (4-s initial IRI). Responding shown on these graphs may be characterized as highly variable both in terms of likelihood of each of the 16 sequences and the lack of any serial dependencies. Figure 10 is the training session event record for subjects in Group 1 (No initial IRI). These graphs display a common pattern of alternating between 2 different sequences. Sequence numbers 6 and 1 are the most common, and they correspond to LLLL and RRRR sequences. Compared to graphs in Figure 9, subjects in the No-IRI condition (Figure 10) show less variability in terms of number of the 16 sequences emitted, show stereotypy in terms of the alternating pattern, and generally take longer to achieve the 100 point deliveries. In summary, the No-IRI subjects show less variability of response sequences.
Figure 9. Training session patterns for subjects in Group 2 (4-s IRI condition). See text for explanation.
Figure 10. Training session patterns for subjects in Group 1 (No-IRI condition). See text for explanation.
Figure 11. Correlograms for subjects in Group 2 (4-s IRI condition). Data is from training session.
Figure 12. Correlograms for subjects in Group 1 (No-IRI condition). Data is from training session.
One way to quantify such variability is by using autocorrelations, which have not been reported in the variability literature. Autocorrelations will detect sequential dependencies that may exist in the data. Unlike the U value, an autocorrelational analysis will indicate degree of dependence between sequences. The autocorrelation function (ACF) correlates all values at time t with all values at time t+k, where k is the lag (the use of the term lag here is to be distinguished from its use in the variability contingency). High correlations, positive or negative, at any lag value indicate serial dependencies in the data; whereas, low correlations indicate a random distribution of the data.

Figures 11 and 12 are training session correlograms for subjects in Groups 2 and 1, respectively. They plot the correlations in the data at lag values 1 to 10. These correlograms show that the subjects in Group 2 have very low correlations indicating a random distribution of sequences through time. Group 1 subjects’ correlograms, on the other hand, show strong serial dependencies in the form of alternations.

Table 3 presents the percent correct and U-values for subjects in the two groups for the training session. Again, as concluded through visual inspection of the event records, subjects in the 4-s IRI condition had higher U-values than subjects in the No-IRI condition. The U-values of the 4-s IRI subjects were 91.75 (CXP), 94.46 (JXP), and 82.34 (MRG). U-values for the other subjects were 24.06 (MXD), 28.90 (AXL), and 49.69 (SDG). To summarize, more differences appear in the variability of the two groups of subjects from the training session.
Table 3

Subjects' percent correct and U-values for training session

<table>
<thead>
<tr>
<th></th>
<th>No-IRI Group</th>
<th>4-s IRI Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>% correct</td>
<td>U-value</td>
<td>% correct</td>
</tr>
<tr>
<td>70.4</td>
<td>24.0</td>
<td>89.3</td>
</tr>
<tr>
<td>96.2</td>
<td>28.9</td>
<td>93.5</td>
</tr>
<tr>
<td>59.5</td>
<td>49.6</td>
<td>94.3</td>
</tr>
</tbody>
</table>

Effects of conditions

The effects the different IRI conditions had on behavior are represented in graphic form by event records and autocorrelations. Figures 13, 14, and 15 are event records for subjects in Group 2. There are three graphs in each figure which represent the last session in each condition for a single subject. For subjects JXP and CXP in Figures 13 and 14, their behavior is highly variable in terms of the bandwidth (emitting all of the 16 sequences with almost equal probability) and in terms of the patterns in which sequences were emitted. Furthermore, both subjects' behavior remains relatively unchanged across conditions. MRG in Figure 15 also starts out highly variable. But exposure to the No-IRI condition reduced the overall variability among the 16 sequences, and it increased stereotypy. The effect of the 8-s condition was a slight increase in overall sequence variability (not as great as the 4-s condition), and a disruption of the stereotypy that developed during the No-IRI condition.

Figures 16, 17, and 18 are the same as 13, 14, and 15 but for Group 1 subjects. The graphs of Group 1 subjects contrast sharply with the Group 2 subjects'
Figure 13. Session patterns of JXP in 4-s IRI condition (upper), No-IRI condition (middle), and 8-s IRI condition (lower). Each graph is the last session of each condition. See text for explanation.
Figure 14. Session patterns of CXP in 4-s IRI condition (upper), No-IRI condition (middle), and 8-s IRI condition (lower). Each graph is the last session of each condition. See text for explanation.
Figure 15. Session patterns of MRG in 4-s IRI condition (upper), No-IRI condition (middle), and 8-s IRI condition (lower). Each graph is the last session of each condition. See text for explanation.
Figure 16. Session patterns of AXL in No IRI condition (upper), 4-s IRI condition (middle), and 8-s IRI condition (lower). Each Graph is the last session of each condition. See text for explanation.
Figure 17. Session patterns of MXD in No IRI condition (upper), 8-s IRI condition (middle), and 4-s IRI condition (lower). Each graph is the last session of each condition. See text for explanation.
Figure 18. Session patterns of SDG in No IRI condition (upper), 8-s IRI condition (middle), and 4-s IRI condition (lower). Each graph is the last session of each condition. See text for explanation.
graphs just discussed. Specifically, the overall bandwidth of the plots is shorter, indicating less variability, than the bandwidths of subjects from Group 2. In addition, the visual pattern of the data indicates response patterning not seen with the previous subjects. For AXL and MXD, there was little change in variability across conditions. For SDG, the No-IRI condition generated a low degree of variability in terms of a small bandwidth and sequence patterning. The effects of the 8-s IRI condition on behavior was that it increased overall sequence variability (bandwidth). It only somewhat disrupted the stereotypic patterns as patterns remain in the middle graph. The return to the No-IRI condition did not recover the original lower degree of overall sequence variability. It did, however, change the patterning from what occurred during the 8-s condition.

Figures 19, 20, and 21 are correlograms for subjects in Group 2. There are three graphs in each figure which represent the last session in each condition. The data are from the same sessions as in Figures 13 through 18. There is some evidence of serial dependencies in these graphs, but degree is slight. Only one data point (Figure 19, first graph at Lag 4) in all of these graphs exceeds 0.5 in either direction. The rest of the points fall around zero indicating no correlation. There is also little evidence of behavior change across conditions. Figure 19 shows that for subject JXP, the 8-s condition produced the closest to random distribution of the data. In Figure 21, the conditions with the IRI produced a slightly more random distribution of the sequences.
Figures 22, 23, and 24 are correlograms for subjects that began in the No-IRI condition. The correlations are much greater for these subjects indicating higher-order response-patterning in the data. Figure 22 (AXL) shows a pattern similar to that shown by the training-session autocorrelational analysis for all Group 1 subjects. The correlations are high (all but 2 are greater than 0.5), and there is little change across conditions. Figure 23 (MXD) also shows high correlations, but there was a change when the 8-s IRI condition was in effect. There was a reduction in the large correlations, particularly at the lower lag values. In essence, the 8-s condition increased this subject’s variability. Another change occurs following a return to the original condition. Note that the upper and lower graphs of Figure 23 are identical, indicating that the subject’s behavior returned to its original state. In fact, inspection of the raw data revealed that MXD emitted sequences in the exactly the same order. Finally, Figure 24 (SDG) shows high correlations in the first two graphs, but these correlations approach zero in the second exposure to the No-IRI condition. This subject’s U-values kept increasing across sessions regardless of condition. Whatever caused this increase, it may be related to the low correlations seen here. A high U-value may be unlikely to appear together with a higher-order pattern because there would be necessarily more sequences to emit in the pattern. A 100 U-value means each of the 16 sequences are equally likely. To have higher-order patterns, subjects would have to emit the 16 different sequences in some order and then repeat the same sequences again. It would be more likely to have higher-order patterning occur with a small number of response possibilities.
The nature of this higher-order response patterning that the event records and the autocorrelations show is interesting. All subjects from Group 1 showed such patterning; that is, they repeated a sequence pattern consistently throughout many sessions. For AXL, the pattern consisted of the eight sequences, RRRR-LLLL-RLLL-LRRR-RRLL-LLRR-RRRL-LLLR, and was produced at various times throughout the experiment. For SDG, a 6-sequence pattern appeared during the 5th session. It was RRRR-LLLL-LLLR-RLRL-LRLR-RRLL-LLRR. This pattern was disrupted during the 8-s IRI condition and never returned. SDG also began emitting a 7-sequence pattern, RRRR-LLLL-RRLR-LLLRLR-RRLL-RRLL-RLRL-LRLR, during session 5. When the condition changed to 8-s IRI, this pattern was slightly disrupted, but it returned to its original state. In the fourth session of this condition responding changed to a 6-sequence pattern. It then reverted back to its original state again when the No-IRI condition was in effect. Only one subject (MRG) from Group 2 showed this kind of patterning, and it did not occur until exposure to the No-IRI condition. The 6-sequence pattern, RRRL-LLLR-RRRR-LLLL-RLRL-RRLL, disappeared when the 8-s IRI condition was reinstated.
Figure 19. Correlograms for subject JXP. Each graph is the last session of each condition: 4-s IRI (upper), No IRI (middle), and 8-s IRI (lower).
Figure 20. Correlograms for subject CXP. Each graph is the last session of each condition: 4-s IRI (upper), No IRI (middle), and 8-s IRI (lower).
Figure 21. Correlograms for subject MRG. Each graph is the last session of each condition: 4-s IRI (upper), No IRI (middle), and 8-s IRI (lower).
Figure 22. Correlograms for subject AXL. Each graph is the last session of each condition: No IRI (upper), 4-s IRI (middle), and 8-s IRI (lower).
Figure 23. Correlograms for subject MXD. Each graph is the last session of each condition: No IRI (upper), 8-s IRI (middle), and No IRI (lower).
Figure 24. Correlograms for subject SDG. Each graph is the last session of each condition: No IRI (upper), 8-s IRI (middle), and No IRI (lower).
CHAPTER IV

DISCUSSION

The results of this experiment show differences in the degree of behavioral variability when subjects were initially exposed to different IRI conditions. This finding is consistent with previous findings in the nonhuman literature (Morris, 1987, 1989; Neuringer, 1992; Page & Neuringer, 1985) that show strong sensitivities to initial IRI conditions. The results, on the other hand, also demonstrate little change in variability when IRIs were manipulated with the different values used in this study. Such lack of differences is not generally consistent with the nonhuman literature (Neuringer, 1991). Finally, the present data suggests that, at least with humans, traditional measures of behavioral variability may need to be supplemented. Discussion is needed regarding the way behavioral variability is defined.

The contrast between variability in nonhuman responding and the results of this study may be seen both in the percent correct and U-values. For example, Neuringer (1991) obtained substantial increases in percent correct and in U-values as IRIs increased. The results of the present experiment, on the other hand, show percent correct to decrease in 5 out of 6 subjects as IRI increases. In addition, U-values remained relatively unchanged across conditions. Only 2 subjects, SDG and MRG, had increasing U-values as IRI increased, and for SDG, it continued to increase when SDG returned to the No-IRI condition. This conclusions reached from
these observations are tenuous given the short sessions and the instability in the data mentioned above. Despite these findings, the differences in the degree of variability observed between the 2 groups, as measured by U-value, autocorrelations, and a visual inspection of the patterns of sequence emission, is interesting and certainly consistent with other findings that show behavioral variability less likely to develop without an IRI between responses (Morris, 1987). The differences in variability between the two groups of subjects are robust and are evident from the first session when they were exposed to the different conditions. Many questions remain, however, about the differences between the humans of the present study and the rats of Neuringer's 1991 study.

For example, why were there differences between human and nonhuman variability when IRI was manipulated? Specifically, why did the subjects' behavior in this study remain so relatively unchanged across conditions? MRG and SDG were the only subjects to show marked change across conditions. These changes only approximated the changes in the behavior of nonhumans. Again, the sessions in this study may have been too short. More sessions should have been run until stability was reached. It seems that when a change did occur across a condition, variability was affected for the remainder of the experiment. SDG's U-value continued to increase even when returned to the No-IRI condition. MRG's U-value was higher in the 4-s IRI than it was in the 8-s IRI, suggesting influence of the No-IRI condition. It therefore seems likely that order effects may occur in human variability. One of the original purposes of this experiment was to examine order effects, but because the
original IRI value (4 s) that was manipulated had insignificant effects on variability, the experiment changed in design and asked additional questions. Further research addressing order effects of IRI manipulations is needed. Perhaps the differences have to do with certain characteristics of human repertoires (e.g., verbal behavior).

Subjects reported remembering series of color combinations (for example, green, green, green, green and then blue, blue, blue, blue) and repeating the combinations in some order, suggesting an intraverbal chain. Continued research is needed to explore this possibility. Small children or persons with retardation may be a place to start. Having the subject engage in some other verbal task or playing background noise also may provide viable avenues of research.

Another difference between the present study and previous studies of variability with nonhumans was the higher-order response patterning observed in the present study. The responding of most subjects showed a patterning that can be described as a stable sequence of 6 or 7 4-response sequences. Such behavior was more likely with the subjects originally exposed to the No-IRI condition, but 1 of the other subjects engaged in this pattern in later No-IRI condition. This phenomenon has not been reported in the nonhuman literature. The difference may represent different characteristics of human and nonhuman responding, or it may be a consequence of measurement practices. The traditional measures of variability do not address the possibility of these higher-order patterns.

Taken alone (or even together with percent correct) the U-value as a measure of variability may be misleading because it cannot give information about the specific
sequences that deviate from random (Machado, 1992). In addition, a high U-value would indicate a higher degree of variability than a low U-value, but there still may be higher-order response patterning in the data. For example, a U-value of 100 would indicate that each of the 16 4-response sequences had an equal probability of occurring; yet this would not preclude higher-order patterns in the data. It would be difficult, but not impossible, to emit all 16 sequences in order and to then continue repeating that order. Such a pattern would produce high autocorrelations. It may be that with a lower U-value, these repetitions become more likely simply because there are fewer members in the higher-order unit. This possibility is supported by the data in the present experiment.

The conclusion here is that U-value and percent correct as measures of behavioral variability may need to be supplemented by other measures. This immediately raises questions of definition. How are we to define behavioral variability and what are the units the variability of which is being measured? The present study has shown two kinds of variability, overall sequence variability and patterns of sequences that occur across time (serial dependency), to be a function of IRI. Autocorrelations have proved useful in the present study in determining the degree of serial dependency in the data. It would be interesting to see what this kind of analysis would reveal in nonhuman studies on behavioral variability. Perhaps the degree of higher-order patterning revealed in the present study is unique to humans.

Despite the questions raised by this experiment, the results contribute to the experimental analysis of operant variability. The experimental analysis of behavioral
variability (Blough, 1966; Cohen, Neuringer, & Rhodes, 1990; Machado, 1989, 1992; Morgan & Neuringer, 1990; Morris, 1987, 1989, 1990; Neuringer, 1991; Page & Neuringer, 1985; Pryor, Haag, & O’Reilly, 1969; Shimp, 1967) has been long overdue. As in evolutionary biology, the mechanisms of variability (which are critical to any selectionist system) were not emphasized early in the science of behavior. This seems to be changing given the increasing number of studies on the topic. The continued analysis of behavioral variability, its mechanisms and its functions, are important to the development of behaviorist theories that view behavior to be the product of both the interaction between selection mechanisms and the level of variability in the current stream of behavior.
REFERENCES


