Distribution Agreement

Kevin Niu March 26th, 2020

A Waiting Game: Concurrent Chains with Equal Primary Reinforcement

by

Kevin Niu

Jack McDowell
Adviser

Neuroscience and Behavioral Biology

Jack McDowell

Adviser

Bob Wyttenbach

Committee Member

Joseph Manns

Committee Member

2020

A Waiting Game: Concurrent Chains with Equal Primary Reinforcement

By

Kevin Niu

Jack McDowell

Adviser

An abstract of
a thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Neuroscience and Behavioral Biology

2020

Abstract

A Waiting Game: Concurrent Chains with Equal Primary Reinforcement

By Kevin Niu

Concurrent chained schedules have long been used to simulate the concept of delayed reinforcement (Chung & Herrnstein, 1967; Fantino, 1969; Squires & Fantino, 1971). Classic models initially developed to predict behavior in these experiments consider the delay period in the context of overall rates of reinforcement to be the key element. These results have been replicated in humans with some discrepancies (Belke, Pierce, & Powell 1989). Older studies tested a very narrow range of interval values and only used scheduled interval values to test their predictions. The current study remedies these issues by including more schedules to improve the testing of models using observed human data. Schedules tested employed a feature used in previous studies, namely, an equal expected time to reinforcement is maintained across alternatives. Four models were tested using data collected from undergraduate psychology students, and were compared using information criteria. Older models that do not include free parameters (Fantino, 1969; Squires & Fantino 1971) accounted for no variation in the data. The data were best described by a modified delay reduction equation featuring a multiplicative relationship between the ratio of reinforcement rate and the ratio of delay reduction, sensitivity parameters for both ratios, and a bias parameter.

Keywords: concurrent chains, delay-reduction theory, delayed reinforcement

A Waiting Game: Concurrent Chains with Equal Primary Reinforcement

By

Kevin Niu

Jack McDowell

Adviser

A thesis submitted to the Faculty of Emory College of Arts and Sciences
of Emory University in partial fulfillment
of the requirements of the degree of
Bachelor of Science with Honors

Neuroscience and Behavioral Biology

2020

Table of Contents

A Waiting Game: Concurrent Chains with Equal Rates of Primary Reinforcement

The fields of science that seek to understand human behavior often overlap in their interests, describing similar phenomena with different terms or studying them with different techniques. Between the psychologists, neuroscientists, economists, and sociologists of history, it’s likely that thousands of years of cumulative time have been spent posing and investigating basic questions of why humans act in the ways that they do. In the past few decades, one topic has risen to prominence among the behavioral sciences; the ways in which humans perceive and interpret the concept of time and how they engage with the future accordingly. For the psychologist, patients all too often engage in self-destructive behaviors in exchange for instant gratification. For the economist, people never seem to save enough money to be financially secure down the road. As humans are constantly required to think about their future and forced to evaluate the long-lasting impacts of their choices, it’s easy to see why this area of research has grown in popularity (Odum, 2011). One integral relationship to understand is the ability of time to influence the value of an event or reward as interpreted in the present. However, in order to understand how people think about the future, we must look back to the past at some of the first landmark experiments that studied the concept of delayed reinforcement.

In order to fully understand previous theories and ideas in this area, we must understand, in detail, the methods used to develop them. Early experiments by behavior analysts to understand response to delayed rewards used pigeons operating on variable interval (VI) schedules (Chung, 1965). In typical experimental designs under VI reinforcement schedules, the pigeon is placed in an operant chamber with a lit key. Time in the schedule is divided into sections or intervals. Any target behavior (e.g., pecking the lit key), that occurs before the time of the interval runs out results in no consequence. A target behavior that occurs after the time of the interval has completed triggers an event, typically the delivery of reinforcement (e.g., presentation of food). Once this occurs, another interval begins. In VI schedules, the amount of time in an interval is randomly selected from an exponential distribution with an experimentally manipulated mean. This allows the environment with which the organism is interacting to be random while still being controlled by the experimenter.

To introduce the concept of choice, behavior analysts use concurrent VI VI schedules (Ferster & Skinner, 1957). In these choice tasks, the operant chamber contains two keys for the pigeon to peck. Each key operates under separate independent VI schedules as described above. Pigeons are free to peck either key at any time and can switch between the keys freely. Choice can be measured as the ratio of pecks that are made to one key over the other or the proportion of pecks made to a single key in comparison to the total pecks observed (Herrnstein, 1961).

An alternate version of VI schedules, known as chained schedules, has been used in past experiments to simulate delayed reinforcement, a topic of interest for this paper (Autor, 1960). In chained schedules, two intervals are joined into a pair: an “initial link” interval that is then followed by a “terminal link” interval. A target behavior completed after the elapse of the initial link interval triggers entry into and initiation of the terminal link interval, accompanied by a stimulus to signal the transition. A target behavior after the elapse of the terminal link interval triggers the delivery of a reward. Just as intervals are randomly generated from an exponential distribution in single-alternative VI schedules, the duration of the initial link and the terminal link can also be randomly derived from separate exponential functions. The mean values for the initial link and terminal link can also be different from one another. Given its association with the reward, the stimulus accompanied by the transition into the terminal link is considered a “conditioned reinforcer.” Whatever stimulus signals the shift from the initial to the terminal link gains reinforcing properties for key pecks because it always precedes the delivery of reinforcement.

In order to apply the same element of choice, two chained schedules can also be presented concurrently. In these concurrently available chained schedules (i.e., “concurrent chains”), there are two sets of initial links and terminal links operating on two separate keys. Both keys are available to the organism for interaction during the initial phase. If an organism enters the terminal link for one of the keys, the other key becomes inactive; typically, this inactivation is indicated by a darkening of the inactive key and produces no consequence for pecking until the next initial link is in effect. Once an organism acquires the scheduled reward during the terminal link, the experiment reenters the initial link phase where both keys become active once again. Figure 1 illustrates visually how the chamber would appear through various stages of the experiment, where dark circles represent inactive keys during the terminal link. Figure 2 presents a chronological map of how the experiment would proceed through time.

Chung (1965) initially used concurrent VI VI schedules to study responses by pigeons in a choice task involving delay to primary reinforcers. This study used concurrent VI60-VI60-s reinforcement schedules. Certain keys would deliver a reinforcement after a blackout period (i.e., both keys become unlit and inoperative for a set amount of time), which simulated a delay. These early investigations showed that the pigeons pecked more on the keys associated with immediate rewards (i.e., shorter delays) than those associated with rewards delivered later (i.e., longer delays). Chung and Herrnstein (1967) sought to determine a more mathematical relationship between temporal elements of the delays and the observed patterns of behavior. The researchers hypothesized a model that states the proportion of behavior allocated to one key could be described as the inverse of the ratio between the delay associated with that key and the sum of the associated delays across both keys:

$$\frac{B_1}{B_1 + B_2} = \frac{t_2}{t_1 + t_2}. \tag{1}$$

where B_1 and B_2 represent the rate of responses on each alternative and t_1 and t_2 represent the duration of the terminal link.

While Equation 1 makes sense intuitively, (i.e., that longer delays on a left key would produce more pecking on a right key), the model is not generalizable to other situations where VI schedules across the alternatives are not the same. Fantino (1969) noted these restrictions and sought to expand upon the previous studies using chained schedules. Fantino (1969) also used pigeons working in operant chambers to collect data, but instead implemented concurrently chains. Across four different conditions, Fantino (1969) maintained the same pair of terminal links, with one key operating on a VI30-s and another on a VI90-s. For three of the conditions, the time of the initial links was kept equal between the alternatives but increased across the conditions, featuring 40-s, 120-s, and 600-s initial links (e.g., a concurrent chain of VI40-s VI30-s vs. VI40-s VI90-s). Fantino (1969) developed a model for his “delay reduction theory,” which predicts that overall reinforcement context (the rate of reinforcement) influences the relationship between delay (the terminal link duration) and allocation of key pecks:

$$\frac{B_1}{B_1 + B_2} = \frac{(T - t_1)}{(T - t_1) + (T - t_2)}. \tag{2}$$

where T represents the overall average time to reinforcement. Fantino (1969) found delay reduction theory supported when he observed the ability of the difference in terminal link rates to produce increased responding to one side break down as the time of the initial link increased. In other words, differences in delays became less noticeable as time spent between earning rewards is longer.

Ignoring the separation of initial and terminal links, concurrent chains are still set to deliver rewards at different rates. Observed differences in response across the keys can simply be explained by differences in the overall rate of reinforcement. In this sense, the model doesn't describe much more than Herrnstein's (1961) original matching law where the proportion of behaviors allocated to a key can be predicted by the proportion of reinforcers obtained from that key.

$$\frac{B_1}{B_1 + B_2} = \frac{R_1}{R_1 + R_2}. \tag{3}$$

Where R_1 and R_2 represent the rate of acquired reinforcement on each alternative. Fantino (1969) recognized this and included a fourth concurrent chain to further test his model. This schedule featured a VI30-s initial link and VI90-s terminal link on one key and a VI90-s initial link and VI30-s terminal link on the other key. Pigeons still adhered well to the delay reduction theory, responding mostly on the VI90-s VI30-s key despite earlier access to entry of the terminal link on the alternative key. Thus, each key is providing food at the same overall rate. However, the pigeons are behaving as if they weren't. While concurrent chains with unequal initial links are not as commonly used in the literature, they illustrate an important idea of how conditioned reinforcement can influence choice.

More recent studies have used similar designs to study how differences in conditioned reinforcement produce preference. McDevitt and Williams (2010) used a schedule like Fantino's (1969) unequal initial link chains in order to study the effects of differential signaling into terminal links. Bell and Williams (2013) used 3-link chained schedules with equal times to reward to further support the notion that elements of conditioned reinforcement can produce biased behavior. Human studies have also incorporated concurrent chained schedules with unequal initial links. Belke, Pierce, and Powell (1989) tested pigeons and human subjects on three of the same schedules used by Fantino (1969) including the schedule with equal rates of primary reinforcement. This human study helps address a concern of concurrent research and behavior analytic research in general: the applicability of theories derived from non-human studies to human behavior. Data obtained from the humans in Belke et al.'s (1989) experiment did not entirely match the performance of the pigeons, inasmuch as the participants' behavior was often less influenced by longer delays. As the researchers theorized, differences in reward types may have contributed to this discrepancy; the humans were compensated with tokens, a reward that is not immediately consumable. Still, human performance in schedules with unequal initial links showed more variability than that of pigeons, suggesting other intrinsic differences between the processes governing decision-making in these two species.

Another hope for future research is the inclusion of a range of schedules with unequal initial link durations. All the previously mentioned studies only included one combination of initial link and terminal link. Just as Fantino (1969) believed Herrnstein's (1961) original model (Equation 1) was only applicable to a specific range of initial and terminal link values, it is possible that the same is true for the delay-reduction theory (Equation 2).

For these reasons, the current project will attempt to better understand the ability of conditioned reinforcement to produce preference and how delays to reinforcement impact choice in humans. To expand upon previous work, this study will include a range of schedules to test multiple initial-link/terminal-link combinations to better understand the interactions between choice and conditioned reinforcement and within-subject consistencies of traditional models.

Method

Participants

Thirty-five undergraduate students were recruited for the study using Emory University's SONA system during the Spring 2020 semester. Students could participate in the study in lieu of a writing assignment for a grade in their introductory psychology courses.

Procedures

Participants underwent a single study visit, held in the principal investigator's (PI) research mentor's laboratory. Participants signed up for a study time slot through the SONA system website with each slot listed as lasting an hour and forty minutes. Before the visit, participants received an email containing the time of their appointment, instructions on how to reach the lab, and a brief overview of the study procedures. Once a participant arrived at the laboratory for their study session, they reviewed and signed an informed consent document with the researcher acting as a witness. Afterwards, the participant was introduced to the computer task. The researcher read a scripted set of instructions on the ways in which the program's environment can be interacted with, and other features of the graphical user interface and their relevance to the experiment (Appendix A). Participants were instructed to earn as many points as possible. Points did not translate into any real rewards for participants, but previous research shows that hypothetical rewards do not illicit behavior significantly different from real rewards (Bickel 2002). Participants were given the opportunity to ask questions before beginning the experiment but answers by the researcher consisted only of repeated information from the scripted instructions. Once the participant had no further questions, the experimenter left the room, and the participant began working on the task. Upon completion of the task, the experimenter re-entered the room to conduct a debriefing session. In this session, the researcher explained the details of the concurrent chains procedures and gave the participant an opportunity to ask questions before being dismissed from the session.

Apparatus

The computer task was administered in a separate closed room. Next to the door was an armchair and against the opposite wall was a desk with a computer apparatus for the experiment to be conducted. The apparatus included a computer monitor, mouse, and keyboard.

Operant Task

The program used to administer the concurrent chained procedure and collect data was coded in VB.net by the PI, using Microsoft Visual Studio 2019. He was assisted by O.L. Calvin and J. J McDowell who, respectively, coded a previous continuous choice procedure and a random number generator class. Both previous programs contributed to certain elements of the code.

Display

The graphical user interface consists of a center button that acts as the target behavior, a changeover key (Findley, 1956), labeled with the text “switch”, which is used to switch between two behavior alternatives, a point counter, a set of schedule-correlated lights, colored boxes associated with reward (green) and punishment (red) signals, and a leaderboard containing top performing participant scores, all of which were arranged in the configuration shown in Figure 3. The red punishment-associated stimulus box, while not utilized in this experiment, was included to maintain the same display of a previous experiments that used a similar procedure to collect human continuous choice data (Klapes, Calvin, & McDowell, 2020)

Gameplay

The experiment begins when the participant presses the spacebar on the keyboard, at which point the center button that reads “START” turns to either blue or yellow with no text. The participant can freely press the space bar in order to earn points, and can press the control button to switch the color of the center button from blue to yellow or from yellow to blue. When either key on the keyboard is pressed, a faint blue border on the corresponding button on the screen will flash to give the appearance of the button physically pressed. Pressing the space bar while the center button is blue and pressing the space bar while the center button is yellow represent two distinct target behaviors linked to separate VI chained schedule sequences. In this phase, the program display is in the configuration as shown in the left panel of Figure 4. If the space bar is pressed after the completion of an initial link criterion, then the terminal link begins. This event is signaled by several discriminative stimuli, including the appearance of an exclamation mark, “!”, on the center button, the removal of the text, “SWITCH”, from the Findley key, and a “blop” sound played from the computer’s speakers. In this phase, the program display is in the configuration as shown in the right panel of Figure 4. At this point, the timer controlling entry into the terminal link on the opposite color is paused and use of the control button to switch between the two colors is disabled. If a target behavior is emitted after the completion of a terminal link criterion, then a reinforcer is delivered, signaled by a flash of the green reward-associated box, an increment of one point on the counter above the center button, and a “ding” sound played from the computer speakers. After the delivery of the reinforcer, the display returns to its initial link configuration. Lastly, a 1-s change-over delay (COD) is implemented, such that any space-bar press made in the initial link that occurs one-second or sooner after a color switch of the center button will produce no consequence regardless of scheduled intervals. The COD is implemented to prevent participants from associating the action of changing over with entry into the terminal link and is commonly used in concurrent schedules (Catania & Cutts, 1963).

Schedules

Seven distinct concurrent chains were used in the experiment. Intervals were generated using a Mersenne Twister 19937 pseudorandom number generator using the values listed in Table 1 as the mean of an exponential distribution. Schedules 1 and 7 were always presented as the first and last schedule respectively while schedules 2-6 were presented in a random order for every participant. Schedule 1 functions as an acquisition stage. Schedule 7 functions as an extinction phase. Boxes shown at the top of the screen display changed throughout the experiment to indicate the schedule that was active with the leftmost box associated with schedule 1 and the rightmost box associated with schedule 7. At the end of each schedule, except for schedule 7 where the experiment ends, the program presents a forced break. A timer counts down on the center button from five-s. At the end of the countdown, a message box appears on the screen with an “okay” button that must be clicked in order to progress to the rest of the experiment. These breaks helped participants further discriminate between schedules and functioned similarly to a brief blackout period. Upon completion of the experiment, if a participant scored within the top 5 scores, they were given the opportunity to enter a three-character expression to represent their score which would be placed on the leaderboard in the display for future participants to see.

Data Collection

The program tracked several measurements of the participants behavior and recorded them as a comma-delimited text file. Tracked information was recorded so that each row in the file contained data for a single initial/terminal link sequence. Collected measurements included the schedule in effect, the number of space bar presses in the initial link on each alternative, the number of space bar presses in the terminal link, the time spent in the initial link, the time spent in terminal link, the number of change overs between the two initial link schedules, and whether the earned reinforcer came from the blue or yellow alternative. Figure 5 shows a sample portion of gameplay from a single initial and terminal link and the accompanying data that would be collected in Table 2.

Analysis

Behavioral Modeling

Four models were tested for each participant. Modeling each participant's behavior using different equations can allow for a comparison to determine the best description of behavior in the face of delayed reinforcement. Two of the tested equations come from previous papers while the other two are newly developed.

Delay Reduction Theory

The first model was Fantino's (1969) original delay reduction theory (Equation 2):

$$\frac{B_1}{B_1+B_2} \begin{cases} = \dfrac{(T-t_1)}{(T-t_1)+(T-t_2)} & \text{when } t_1 < T \; and \; t_2 < T \\ = 1 & \text{when } t_2 > T \\ = 0 & \text{when } t_1 > T \end{cases} \qquad (2)$$

Recall that B_1 and B_2 denote rates of responding on each key, T denotes the average time to reinforcement in the environment, and t_1 and t_2 represent the average duration of the terminal links in each component. Note that the original equation does include cases of exclusive preference when a terminal link duration is greater than the average rate of reinforcement. These conditions for exclusive preference were applied to all the models discussed henceforth.

A Modified Delay Reduction Theory

$$\frac{B_1}{B_1+B_2} = \frac{R_1(T-t_1)}{R_1(T-t_1)+R_2(T-t_2)} \tag{4}$$

The second model is from Squires and Fantino (1979). This follow-up to Fantino (1969) introduced the rate of reinforcement as an additional measure to predict choice behavior. This model is the same as Fantino's (1969) original model (Equation 2) with the inclusion of R_1 and R_2 as the rates of acquired primary reinforcement on each key. These factors were added to better account for concurrent chained schedules with different initial link values. In the original equation, two concurrent chained schedules with equal valued terminal links should produce a 50-50 allocation of behavior, regardless of differences in initial link and rates of primary reinforcement. Including rates of reinforcement in the model captures the effect of the initial link. This new model carries another important feature; the reversion to the matching law (Equation 3) when terminal link values are zero, an environment that is essentially just a concurrent VI VI choice task. This equation is a significant improvement over Fantino's (1969) original equation (Equation 2) now that the model is generalizable to other concurrent VI VI environments and not just chained schedules.

A Generalized Delay Reduction Theory

The third model incorporates features from the generalized matching law (GML) and Fantino's (1969) original delay reduction theory. The GML (Baum, 1974),

$$\frac{B_1}{B_2} = b\left(\frac{R_1}{R_2}\right)^a \tag{5}$$

is an iteration of the matching law that adds to Herrnstein's (1961) original equation by introducing two free parameters. The first parameter, a, represents sensitivity to rate of reinforcement, with 1 being a perfect matching value and 0 representing no sensitivity to the reinforcers. The second parameter, *b*, represents a bias value that captures any other systematic differences between the two options that isn't captured by a difference in the acquired reinforcement rate. These two parameters of bias and sensitivity can be applied to the delay reduction as:

$$\frac{B_1}{B_2} = b\left(\frac{T - t_1}{T - t_2}\right)^{a} \qquad (6)$$

with *a* now describing sensitivity to differences in delay reduction as opposed to differences in acquired reinforcement.

Note that the original delay reduction equation (Equation 2) is algebraically transformed from a proportion to a ratio version for Equation 6. This transformation is to mirror the structure of the GML so that additional parameters from this model can be included for the second transformation, the power-function inclusion of the bias and sensitivity parameter. This new model will allow for individual differences between participants and does not predict identical behavior across organisms for a single schedule.

A Modified Generalized Delay Reduction Theory

$$\frac{B_1}{B_2} = b\left(\frac{R_1}{R_2}\right)^{a_R}\left(\frac{T - t_1}{T - t_2}\right)^{a_T} \qquad (7)$$

The final model that was tested applies the ratio transformation and the addition of free parameters of the generalized matching law to the modified delay reduction theory from Squires and Fantino (1971; Equation 4). However, this equation contains separate sensitivity parameters for the rate of reinforcement and delay reduction. This equation maintains the property of collapsing to the GML when terminal link values are zero, just as Equation 4 collapses to Hernstein's (1961) original matching law when terminal link values are zero.

Preprocessing

An important difference between the current project and previous work is the use of observed data to generate predictions of models. While older studies used *scheduled* delays, or simply the mean of the exponential distribution from which intervals were derived, the operant task used in this experiment captures *observed* delays from every reinforcer acquired, which should provide a more accurate description of the environment encountered by participants.

When creating and evaluating models for behavior in these experiments involving schedules of reinforcement, each schedule is represented by a single data point. For each schedule across participants, the total number of initial link presses on the blue alternative, the total number of initial link presses on the yellow alternative, number of reinforcers acquired on the blue alternative, and reinforcers acquired on the yellow alternative were summed. These values were used to calculate proportion of presses allocated to each side and the overall rate of reinforcement for each alternative. Durations of the all initial link and terminal link within a schedule were averaged.

Observed data for each participant were obtained using the number of space bar presses allocated to each alternative, calculated as either a proportion or a ratio depending on the model tested. Predicted data were be calculated according to the four models discussed above using data collected by the program (viz., Equations 2,4,6, and 7). The number of reinforcers acquired on each side was used to calculate the rate of reinforcement. The sum of initial and terminal link durations was used to calculate the average time to reinforcement. It is important to note that while the models referred to previously are all represented in their original forms, the actual analysis was carried out using log-transformed, ratio versions of all the models (Appendix B). Logarithmic transformations of the GML allow the model to conform to a line and facilitate the estimation of parameters where *a* and *b* become the slope and intercept respectively (Davison & Tustin, 1978). The ratio transformation of Equations 2 and 4 were carried out to maintain consistency with the format of the GML.

Several steps were taken to identify viable data for modeling. These measures were taken to ensure that analyzed data represented active and engaged behavior by participants that reflected changes in the primary and conditioned reinforcement cues. First, test data from the first four participants were not included in the analyses. Significant changes were made to the program to improve its performance after the first four participants. For the remaining participants, data from the first five reinforcers of a schedule were eliminated for all schedules and all participants. In these early stages of a schedule, participants had very little exposure and experience with the options provided. Observed behavior likely reflected an exploratory and transitional period and is unlikely to be representative of steady-state equilibrium behavior that the models of interest are meant to describe. Additionally, if the first five reinforcers of a schedule were all acquired from the same color, data from subsequent reinforcers would also be discarded until at least one reinforcer was gained on the opposite color. This was done for similar reasons, as participants could not clearly evaluate their environment and allocate behavior accordingly without having experience from both choices. For three participants, exclusive preference conditions were met, where the average acquired length of terminal link duration on one alternative was longer than the overall average time to reinforcement, as noted in Equation 2. Unfortunately, this meant that data from these participants was unanalyzable due to logarithmic transformation. For another three participants, responding occurred on only one alternative. Data from these participants were also discarded as they would be unanalyzable in the context of delay reduction given that no delay was experienced on one of the alternatives. The resulting pool of data narrowed to 25 participants.

Results

For an initial analysis, the proportion of variance accounted for (PVAF) by each model was calculated across all participants and is listed in the fourth column of Table 3. While this provides useful individual information on the predictive power of each model, it does not necessarily allow us to fairly compare across them. The generalized models (Equations 6 and 7) both contain parameters that can change between participants. Of course, these models will explain more variance given their increased complexity. For this reason, information criteria were used to calculate an estimator that would allow for comparison across models with differing levels of complexity (Akaike, 1992). Two information criterion estimators were used: the Akaike Information Criterion corrected (AICc), which is given by

$$AICc = n\,ln\left(\frac{RSS}{n}\right) + 2K\left(\frac{n}{n-K-1}\right),$$

-and Bayesian Information Criterion (BIC), given by

$$BIC = n\,ln\left(\frac{RSS}{n}\right) + Kl\,n(n),$$

where *n* is the number of data points, *RSS* is the residual sum of squares, and *K* is the number of free parameters.

For both AICc and BIC, a smaller value indicates a more robust model. A smaller residual sum of squares and fewer parameters are the combination favored by these equations. By comparing AICc and BIC values, the “best” model within the set tested was determined for each participant. The number of participants for which each model was deemed superior for each information criterion across each model is shown in Figure 6. By this measure, the AICc slightly favored Equation 6 while the BIC strongly favors Equation 7. Another way to compare the merits of each model is to use an aggregate information criterion estimator value. McArdle, Davison, and Navakatikyan (2007) devised a way to calculate an information criterion value across multiple data sets:

$$AICc = \sum_{i=1}^{N}[n_i * ln\,(\frac{RSS}{n})] + 2K\sum_{i=1}^{N}\left(\frac{n_i}{n_i - K - 1}\right),$$

and

$$BIC = \sum_{i=1}^{N}[n_i * ln\,(\frac{RSS}{n})] + KNl\,n(n_t).$$

The quantities in these equations are the same as in the standard AICc and BIC calculations for a single data set. For the aggregate estimator, *N* is the number of data sets, *i* is the index number of the data set, and n_t is the total number of data points across all the data sets. These aggregate values, shown in the last two columns of Table 3, can provide a more complete picture of the strength of each model across the entire data set. Lastly, average parameter estimates for Equations 6 and 7 are shown in Table 4.

In terms of VAF, Equation 7 was the clear winner (Table 3). Equations 2 and 4 actually accounted for no variance in the observed data. For a majority of the participants, these models were no better at explaning the behavior of the participants than the simple mean of the observed behavior. However, when taking into account the number of parameters in each model, Equation 7 is heavily penalized by the AICc and to a lesser degree by the BIC. For the AICc estimator, Equation 6 is the superior model, supported by the fact that this model had the best individual AICc value for the most number of participants and produced the best aggregate value. When examining the BIC estimator, the favored model is not as clear. According to the aggregate calculation across all of the data sets, Equation 4, the modified delay reduction equation with no sensitivity or bias parameters, does the best. However, Equation 7 produced the best individual model for the greatest number of participants. Finally, when examining the parameters of the generalized models, the sensitivity paramemters of both delay reduction and reinforcement both indicate undermatching. The bias parameter also suggests that there was little to no systematic bias between the blue and yellow alternatives.

Discussion

There are many considerations when attempting to indentify a best model based on the results of the statistical analysis, especially when these results indicate multiple possibilities. The first item to consider is whether conclusions should rely more heavily on the AICc or BIC calculations. Most model selection papers in the literature present both information criterion values. If they both indicate the same best model, then the interpretation is rather simple. However, in this instance, the two measures do not agree. Previous papers have only included the BIC where there was limited amount of data within a data set (Navakatikyan & Davison, 2010). In the case of complex models, AICc is extremely conservative, especially when *n*, the number of data points, is small. While experts do note that AICc is typically used for smaller sample sizes, such as the data sets used in this experiment, they also explain that AICc chooses a model for that specific sample size, while BIC assumes a true best model independent of *n* (Burnham & Anderson, 2004). Given the number of data points for each participant is only six, in comparison to Equation 7 with three free parameters, and that we hope to find a model generalizable to situations outside of the given context, BIC may be a more appropriate tool in helping us understand our results.

Based on the aggregate BIC values alone listed in Table 3, the modified generalized delay reduction theory (Squires and Fantino, 1971; Equation 4) serves as the best of the four models for behavior in a concurrent chains situation. This would be an easy conclusion if not for two details that are difficult to overlook. First, the generalized version of Equation 4 (Equation 7) produced the best BIC value for most participants. Second, Equation 4 explains no variance of the data for so many of the participants. To discuss the implications of a model that accounted for no variance is not a great use of time. For this reason, Equations 2 and 4 will not be considered. With the remaining choices, Equation 7 provided a better fit than Equation 6, for the aggregate BIC value and 16 of 25 individual values. Further supporting this notion is the amount of variance accounted for by each of these equations, with Equation 7 accounting for almost three-quarters of the variance in the data while Equation 6 accounted for less than half. This information supports the argument for the inclusion of sensitivity and bias parameters as well as the rate of reinforcement for any predictive power of delay reduction.

While it may be a more complex equation, the generalized version of the modified delay reduction theory (Equation 7) does have a strong theoretical basis. The GML has long been regarded as a powerful description of behavior in VI VI schedule experiments. It has even been shown to apply just as well to non-engineered natural environments (McDowell & Caron 2010; Houston, 1986). The fact that Equation 7 reduces to the GML when terminal link values are zero is an advantage it holds over the other models considered in this paper. The sound basis of the GML makes Equation 7 a proper description of behavior for situations outside of concurrent chained schedules. Equation 7 also reflects other versions of the GML. The matching law has been extended to a variety of experimental designs. One specific version assumes that other attributes of a reinforcer, outside of its rate of presentation, can be included in the GML in a multiplicative relationship with reinforcement rate (Baum & Rachlin, 1969). Known as the concatenated generalized matching law (cGML), this equation is most often used to study reinforcers that vary by magnitude (Cording, McLean, & Grace, 2011; Davison & McCarthy, 1988).

$$\frac{B_1}{B_2} = b\left(\frac{R_1}{R_2}\right)^{a_R}\left(\frac{M_1}{M_2}\right)^{a_M}. \tag{8}$$

M_1 and M_2 represent the magnitude of the two reinforcers obtained from each alternative and a_M represents the sensitivity to the magnitude of the reinforcement, separate from its rate. Equation 7 proposed in this paper is very similar to the cGML, only substituting the ratio of delay reduction for the ratio of reinforcer magnitude. From this perspective, delay reduction, or possibly the concept of immediacy, can be considered a feature of the reinforcer that changes its effective value (Rachlin, 1971). Given experimental evidence from individual BIC measures, the amount of variance explained, and its theoretical foundations, Equation 7 is the most appropriate model for behavior in a concurrent chains schedule design.

Still, given certain outcomes of the current study, future investigation into these models is necessary. One notable result is the failure of the non-generalized equations to describe any of the participants' behavior. While few studies of concurrent chains in humans have attempted to conduct the in-depth analysis in the context of delay-reduction presented here, Belke et al.'s (1989) study did replicate Fantino's (1969) original study with human subjects. While Belke et al.'s experiment did observe deviations from values predicted by both the original model and the modified version, those deviations were not as extreme as those observed in this study.

One might consider the obvious biological difference between humans and pigeons as an explanation for why Fantino's (1969) theories worked better for the birds' data in his original studies. Even Belke et al.'s (1989) study found that humans followed more of an optimization model in some of the concurrent chained schedules to which they were exposed. However, there are procedural differences that may confound these conclusions. In general, pigeons have been found to exhibit the same type of hyperbolic discounting as humans (Vanderveldt, Oliviera, & Green, 2016; Calvert, Green, & Myerson, 2011). Neurological studies have supported the notion that mammals and birds maintain temporal and cost-based discounting function in brain regions like the ventral striatum, pre-frontal cortex and their avian analogs (Izawa, Zachar, Yanagihara, and Matsushima, 2003; Dykes, Porter, and Colombo, 2019; Wittmann, Lovero, Lane, and Paulus, 2010). Given the universal importance of decision-making processes that can evaluate the expected value of outcomes in the face different attributes like probability, delay, and magnitude, it's very possible that the evolutionarily conserved functions of these regions should produce similar behavior under such circumstances. For these reasons, it may be necessary to entertain procedural and analytical differences that could contribute to the observed difference in human performance in the current study.

In comparison to previous studies, there are several experimental and analytical aspects that could have resulted in the observed disparity. The in-depth analysis of the current study may contribute to this apparent difference. While Belke et al.'s (1989) study only included three schedules and Fantino's (1969) original study only included four, the current study utilized six schedules to produce the estimations of each of the considered models. The previous studies only calculated individual deviations for single schedules independently. As mentioned previously, past studies also only used scheduled, as opposed to observed, delays to produce model predictions. Additionally, previous human studies present participants with tokens as reinforcers that can be exchanged for real money after the end of a session. Comparisons of studies that use real and hypothetical monetary rewards do not show significant differences in the context of delayed reinforcement (Johnson and Bickel, 2002). However, monetary rewards themselves are not immediately consumable reinforcement like food is for pigeons, especially when received only at the end of a session. Differences in reinforcer consumption, value, and delivery could explain the lack of continuity with Belke et al.'s(1989) previous assessment.

These discussed elements of past and current experiments can inform directions for future studies. The conflicting conclusions reached by different metrics of the information criterion values used suggests that future studies should include schedules to increase the size of each data set. The increased statistical power afforded by more schedules may hopefully counter the harsh effects of the AICc, allowing for a conservative yet balanced assessment of generalized models. The separation of sensitivity parameters for delay and reinforcement is another novel aspect of the current analysis that holds potential in the future. A 2013 study used the sensitivity parameter of the generalized matching law to quantify behavioral effects of medication for adolescents with ADHD (Yan-Ling & Xu, 2013). The derivation of a sensitivity to delay reduction parameter can allow for another way to quantify changes in impulsivity, self-control, or temporal discounting in response to behavioral or pharmacological interventions. Whatever the future of delayed reinforcement research holds, it's evident that individuals relate to their futures in complex and unique ways. While these inconsistencies may frustrate the researchers that seek to describe them, they no doubt contribute to a vast diversity of character and value in our species.

References

Akaike, H. (1992). Information theory and an extension of the maximum likelihood principle. *Springer Series in Statistics Breakthroughs in Statistics*, 610–624. doi: 10.1007/978-1-4612-0919-5_38

Autor, S. M. (1960). The strength of conditioned reinforcers as a function of frequency and probability of reinforcement. Unpublished doctoral dissertation. Harvard University

Baum, W. M. (1974). On two types of deviation from the matching law: bias and undermatching. *Journal of the Experimental Analysis of Behavior*, *22*, 231–242. doi: 10.1901/jeab.1974.22-231

Baum, W. M., & Rachlin, H. C. (1969). Choice as time allocation1. *Journal of the Experimental Analysis of Behavior*, *12*, 861–874. doi: 10.1901/jeab.1969.12-861

Belke, T. W., Pierce, W. D., & Powell, R. A. (1989). Determinants of choice for pigeons and humans on concurrent-chains schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, *52*, 97–109. doi: 10.1901/jeab.1989.52-97

Bell, M. C., & Williams, B. A. (2013). Conditioned reinforcement in chain schedules when time to reinforcement is held Constant. *Journal of the Experimental Analysis of Behavior*, *99, 179–188. doi: 10.1002/jeab.10*

Burnham, K. P., & Anderson, D. R. (2004). Multimodel Inference. *Sociological Methods & Research*, *33*, 261–304. doi: 10.1177/0049124104268644

Calvert, A. L., Green, L., & Myerson, J. (2011). Discounting in pigeons when the choice is between two delayed rewards: implications for species comparisons. *Frontiers in Neuroscience*, *5*. doi: 10.3389/fnins.2011.00096

Catania, A. C., & Cutts, D. (1963). Experimental control of superstitious responding in humans. *Journal of the Experimental Analysis of Behavior*, *6*, 203-208. doi:10.1901/jeab.1963.6-203

Chung, S.-H., & Herrnstein, R. J. (1967). Choice and delay of reinforcement1. *Journal of the Experimental Analysis of Behavior*, *10*, 67–74. doi: 10.1901/jeab.1967.10-67

Chung, S.-H. (1965). Effects of delayed reinforcement in a concurrent situation. *Journal of the Experimental Analysis of Behavior*, *8*, 439–444. doi: 10.1901/jeab.1965.8-439

Cording, J. R., Mclean, A. P., & Grace, R. C. (2011). Testing the linearity and independence assumptions of the generalized matching law for reinforcer magnitude: A residual meta-analysis. *Behavioural Processes*, *87*, 64–70. doi: 10.1016/j.beproc.2011.02.011

Davison, M. C., & Tustin, R. D. (1978). The relation between the generalized matching law and signal-detection theory. *Journal of the Experimental Analysis of Behavior*, *29*, 331–336. doi: 10.1901/jeab.1978.29-331

Davison, M., & McCarthy, D. (1988). *The matching law: A research review*. Hillsdale, NJ: Erlbaum.

Dykes, M., Porter, B., & Colombo, M. (2019). Neurons in the pigeon nidopallium caudolaterale, but not the corticoidea dorsolateralis, display value and effort discounting activity. *Scientific Reports*, *9*. doi: 10.1038/s41598-019-52216-3

Fantino, E. C. (1969). Choice and rate of reinforcement. *Journal of the Experimental Analysis of Behavior*, *12*, 723-730.

Ferster, C.B., & Skinner, B. F. (1957). Schedules of reinforcement. *Appleton-Century-Crofts*. https://doi.org/10.1037/10627-000

Findley, J. D. (1958). Preference and switching under concurrent scheduling. *Journal of the Experimental Analysis of Behavior, 1*, 123-144. doi: 10.1901/jeab.1958.1-123

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement1, 2. *Journal of the Experimental Analysis of Behavior, 4*, 267–272. doi: 10.1901/jeab.1961.4-267

Houston, A. (1986). The matching law applies To wagtails foraging in the wild. *Journal of the Experimental Analysis of Behavior, 45*, 15–18. doi: 10.1901/jeab.1986.45-15

Izawa, E.-I., Zachar, G., Yanagihara, S., & Matsushima, T. (2003). Localized lesion of caudal part of lobus parolfactorius caused impulsive choice in the domestic chick: evolutionarily conserved function of ventral striatum. *The Journal of Neuroscience, 23*, 1894–1902. doi: 10.1523/jneurosci.23-05-01894.2003

Johnson, M. W., & Bickel, W. K. (2002). Within-subject comparison of real and hypothetical money rewards in delay discounting. *Journal of the Experimental Analysis of Behavior, 77* , 129–146. doi: 10.1901/jeab.2002.77-129

Klapes, B., Calvin, O. L., & McDowell, J. J. (2020). A discriminated rapid-acquisition laboratory procedure for human continuous choice. Manuscript in preparation.

McArdle, B., Navakatikyan, M. A., & Davison, M. (2007). Application of information criteria to behavioral studies. Unpublished manuscript, Departments of Statistics and Psychology, University of Auckland, Auckland, New Zealand

McDevitt, M. A., & Williams, B. A. (2010). Dual effects on choice of conditioned reinforcement frequency and conditioned reinforcement value. *Journal of the Experimental Analysis of Behavior, 93*, 147–155. doi: 10.1901/jeab.2010.93-147

McDowell, J. J., & Caron, M. L. (2010a). Matching in an undisturbed natural human environment. *Journal of the Experimental Analysis of Behavior*, *93*, 413-431. doi: 10.1901/jeab.2010.93-415

Navakatikyan, M. A., & Davison, M. (2010). The dynamics of the law of effect: a comparison of Models. *Journal of the Experimental Analysis of Behavior*, *93*, 91–127. doi: 10.1901/jeab.2010.93-91

Odum, A. L. (2011). Delay Discounting: Im a k, you're a k. *Journal of the Experimental Analysis of Behavior*, *96*, 427–439. doi: 10.1901/jeab.2011.96-423

Rachlin, H. (1971). On the tautology of the matching Law. *Journal of the Experimental Analysis of Behavior*, *15*, 249–251. doi: 10.1901/jeab.1971.15-249

Squires, N., & Fantino, E. (1971). A model for choice in simple concurrent and concurrent-chains schedules. *Journal of the Experimental Analysis of Behavior*, *15*, 27–38. doi: 10.1901/jeab.1971.15-27

Vanderveldt, A., Oliveira, L., & Green, L. (2016). Delay discounting: pigeon, rat, human—does it matter? *Journal of Experimental Psychology: Animal Learning and Cognition*, *42*, 141–162. doi: 10.1037/xan0000097

Wittmann, M., Lovero, K. L., Lane, S. D., & Paulus, M. P. (2010). Now or later? Striatum and insula activation to immediate versus delayed rewards. *Journal of Neuroscience, Psychology, and Economics*, *3*, 15–26. doi: 10.1037/a0017252

Yan-Ling, R., & Xu, D. (2013). Effects of methylphenidate in children with attention deficit hyperactivity disorder: A comparison of behavioral results and event–related potentials. *Attention Deficit Hyperactivity Disorder in Children and Adolescents*. doi: 10.5772/53782

Appendices

A: Script

“Before we begin, I would like to go over the Informed Consent with you. This document describes your rights and privileges as a participant in the study, if you choose to participate.”

“The current study aims to understand how people behave in changing environments with a focus on temporal influences. The experiment will involve a brief demographic survey and a computer task. The computer task will ask you to interact with a digital environment by pressing buttons and features flashing lights and sounds. There are no foreseeable risks to the study. At the same time, there are no foreseeable benefits to you as an individual outside of the credit you are receiving for your class. You may choose to drop out of the study at any time, just let the research staff know. If you agree to these aspects of the study and to participate in the experiment, you may sign the informed consent form.

“Would you like a copy of this document for your records?”

> ***If affirmative:*** *while participant is performing the study, make a copy of the informed consent sheet on the 4th floor. Give the copy it to participant during the debriefing period.*
>
> ***If negative:*** *Move on.*

“Next, I’m going to ask you to fill out this quick demographic information sheet. Information on this sheet will be deidentified from your name and stored only by your participant ID in our database.”

Set up protocol, inputting the participants three digit number in the ID section

“As discussed in the informed consent, thank you again for choosing to participate in this study. This task investigates how people behave in changing environments. Your goal is to earn as many points as possible. You can earn points by pressing the ‘space’ bar on the keyboard. The center button that says ‘START’ on it is currently grey, but it will change to either blue or yellow when the experiment begins. You can switch between blue or yellow by pressing the ‘ctrl’ button on the keyboard. At certain points during the experiment, an exclamation mark will appear on the center button accompanied by a “blop” sound. During this time, you will not be able to switch between blue and yellow but you may continue pressing space bar to receive points. The rate at which you can gain points is different depending on whether this button is blue or yellow. However, both colors will always provide opportunities to gain points for pressing the space bar.”

“Your mission is to figure out how to earn as many points as possible. Visual and auditory cues will help you with this task. Every time that you earn a point, the green light to the left of the center button will flash and a ‘ding’ will sound. You can see that there are also lights at the top of the screen. As these lights change, the effects of blue and yellow also change.”

“Throughout the experiment, you will receive forced 5second breaks. You may take longer than the allotted 5 seconds if necessary. You may exit the laboratory if you wish; there is water, coffee and tea to the right and a restroom to your left at the end of the hall on the left-hand side. When you are finished taking your break, use the trackpad to left-click ‘Ok’ on the message prompt, and then left-click again on the screen to begin acquiring points again.”

“Lastly, you’ll notice a leaderboard on the right-hand side of the screen. If you score enough points, then a three-character expression of your choice will be placed on this board for other participants to see. Once you begin, focus on the task and refrain from other activities, such as using your phone.”

B: Transformed Equations

Model	Eq.	Log/Ratio Transformation
$\frac{B_1}{B_1+B_2} = \frac{(T-t_1)}{(T-t_1)+(T-t_2)}$	2	$\log\left(\frac{B_1}{B_2}\right) = \log\left(\frac{(T-t_1)}{(T-t_2)}\right)$
$\frac{B_1}{B_1+B_2} = \frac{R_1(T-t_1)}{R_1(T-t_1)+R_2(T-t_2)}$	4	$\log\left(\frac{B_1}{B_2}\right) = \log\left(\frac{R_1}{R_2}\right) + \log\left(\frac{(T-t_1)}{(T-t_2)}\right)$
$\frac{B_1}{B_2} = b\left(\frac{T-t_1}{T-t_2}\right)^a$	6	$\log\left(\frac{B_1}{B_2}\right) = a\log\left(\frac{(T-t_1)}{(T-t_2)}\right) + \log(b)$
$\frac{B_1}{B_2} = b\left(\frac{R_1}{R_2}\right)^{a_R}\left(\frac{T-t_1}{T-t_2}\right)^{a_T}$	7	$\log\left(\frac{B_1}{B_2}\right) = a_R\log\left(\frac{R_1}{R_2}\right) + a_T\log\left(\frac{(T-t_1)}{(T-t_2)}\right) + \log(b)$

Tables

Table 1

Program Schedules

		Interval Values (s)			
Schedule #	(B) initial	(B) terminal	(Y) initial	(Y) terminal	Exposure(s)
1	1	1	1	2	420
2	4	8	8	4	840
3	5	7	7	5	840
4	6	6	6	6	840
5	7	5	5	7	840
6	8	4	4	8	840
7	9	9	9	9	180

Note. (B) = blue alternative, (Y) = yellow alternative

Table 2

Example Gameplay

Schedule	# of Blue Presses Initial Link	# of Yellow Presses Initial Link	# of Presses Terminal Link	Blue Reinforcer	Yellow Reinforcer	Initial Link Delay	Terminal Link Delay
1	7	11	6	0	1	7.32	3.14

Table 3

Information Criterion Values

Model	Equation	Name	Average PVAF	Aggregate AICc value	Aggregate BIC value
$\frac{B_1}{B_1 + B_2} = \frac{(T - t_1)}{(T - t_1) + (T - t_2)}$	2	Delay Reduction	0	-378.7	-378.7
$\frac{B_1}{B_1 + B_2} = \frac{R_1(T - t_1)}{R_1(T - t_1) + R_2(T - t_2)}$	4	Modified Delay Reduction	0	-389.3	-389.3
$\frac{B_1}{B_2} = b\left(\frac{T - t_1}{T - t_2}\right)^{a}$	6	Generalized Delay Reduction	0.427	-402.5	-352.0
$\frac{B_1}{B_2} = b\left(\frac{R_1}{R_2}\right)^{a_R}\left(\frac{T - t_1}{T - t_2}\right)^{a_T}$	7	Modified Generalized Delay Reduction	0.723	-295.4	-369.7

Table 4

Parameter Values

Model	Equation	Name	Average Parameter Values
$\frac{B_1}{B_2} = b\left(\frac{T - t_1}{T - t_2}\right)^{a}$	6	Generalized Delay Reduction	a = .370 b = 1.095
$\frac{B_1}{B_2} = b\left(\frac{R_1}{R_2}\right)^{a_R}\left(\frac{T - t_1}{T - t_2}\right)^{a_T}$	7	Modified Generalized Delay Reduction	a_R = .481 a_T = .409 b = 1.051

Figure Captions

Figure 1. Concurrent chained schedules procedure illustrating possible pathways. Dark circles represent inactive keys

Figure 2. Concurrent chained procedure illustrating scheduled intervals across alternatives

Figure 3. Starting screen of task

Figure 4. The graphical user interface that accompanies each phase of gameplay

Figure 5. an example of one initial and terminal link sequence of gameplay

Figure 6. Each bar represents the number of participants for which that model was deemed the “best” for each IC estimator

Figures

Figure 1. Niu

Concurrent Chains Keys

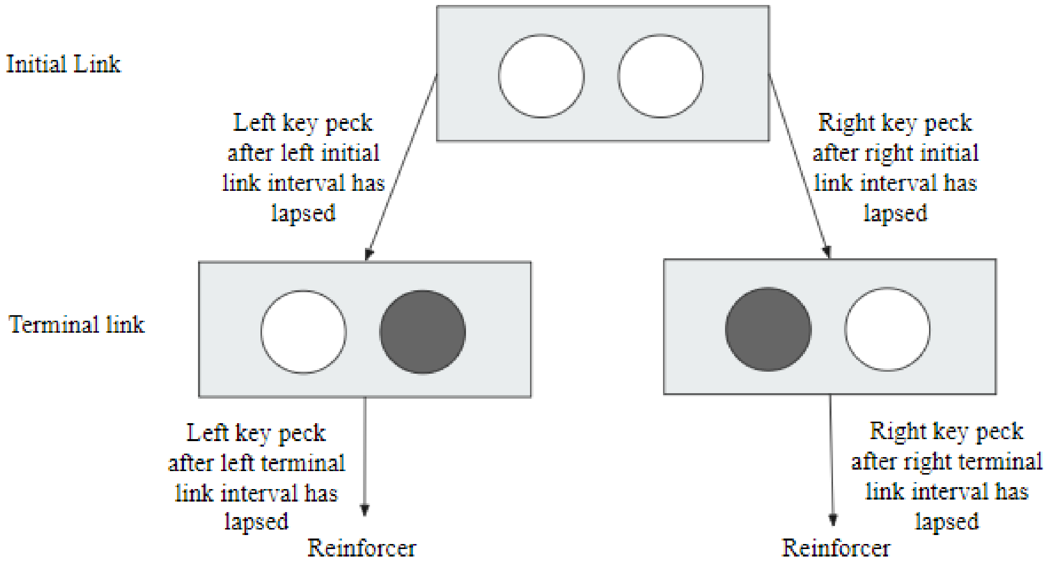


Figure 2. Niu

Concurrent Chains Intervals

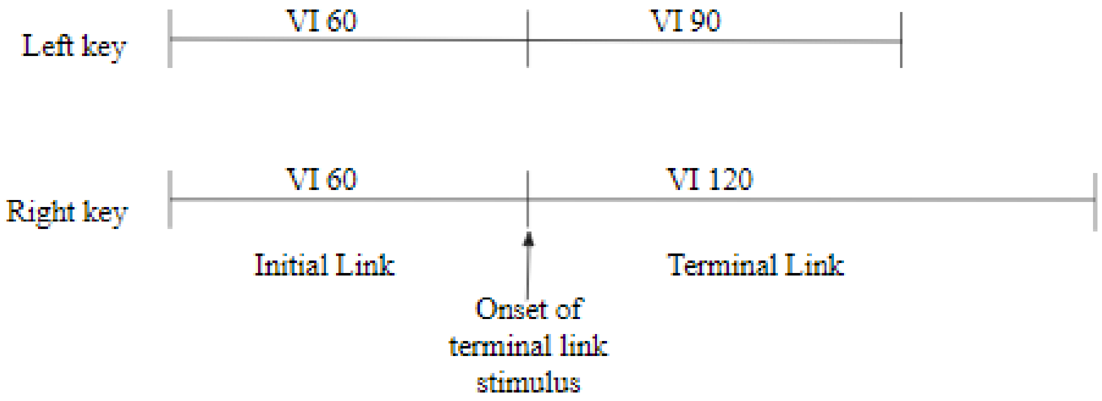
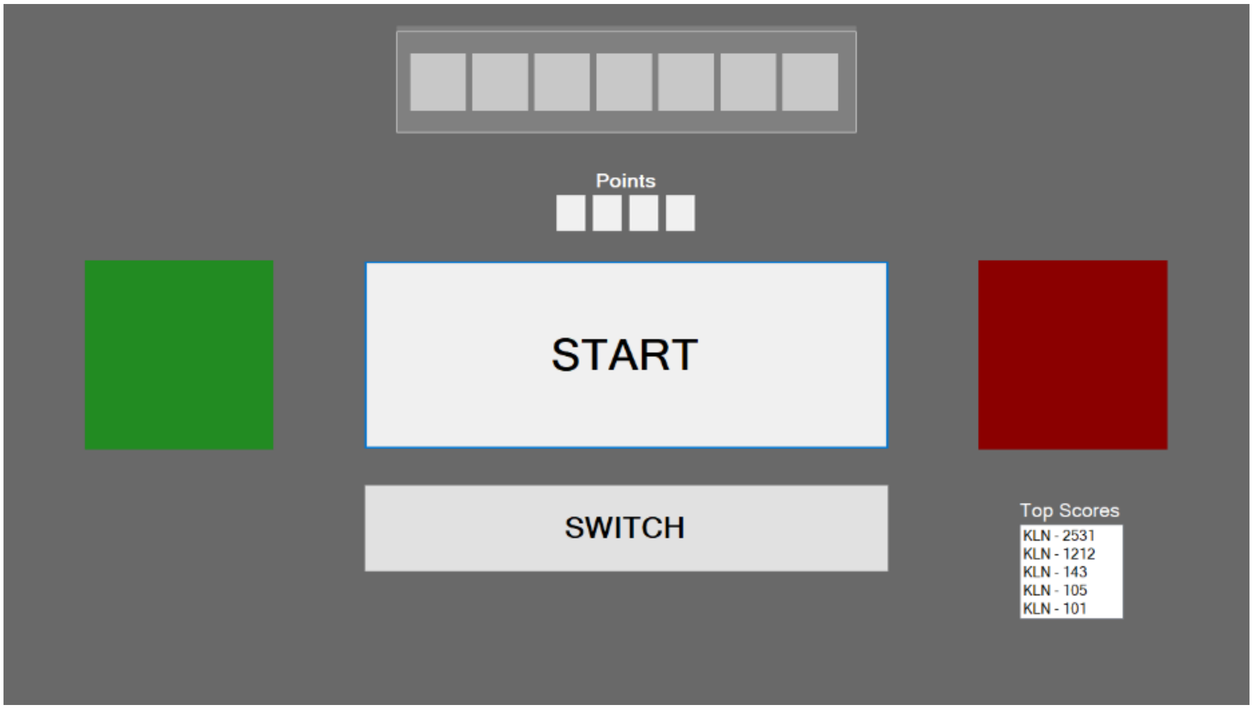


Figure 3. Niu

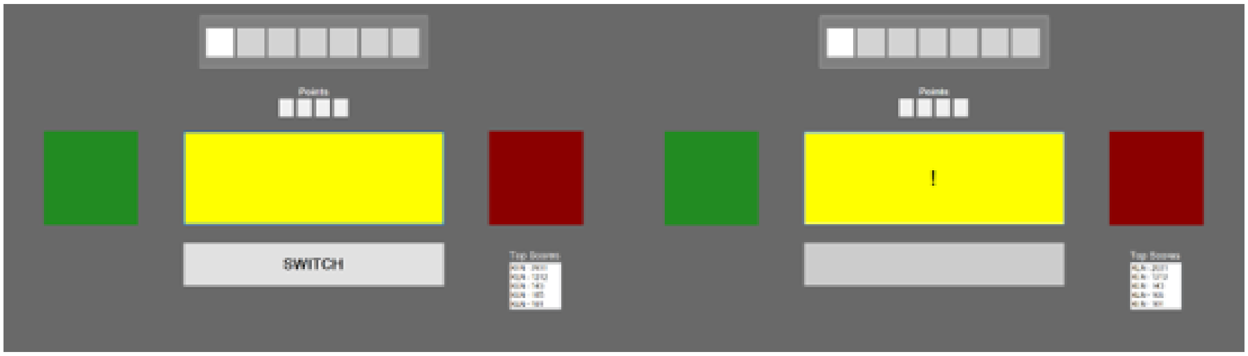
Starting GUI



KLN - 2531
KLN - 1212
KLN - 143
KLN - 105
KLN - 101

Figure 4. Niu

Initial and Terminal Link Display



Initial Link Configuration

Terminal Link Configuration

-Findley key enabled, participant can freely switch between blue and yellow -Center button has no text	-Findley key disabled, participant cannot switch between blue and yellow -Center button displays “!”

Figure 5. Niu

Example gameplay sequence

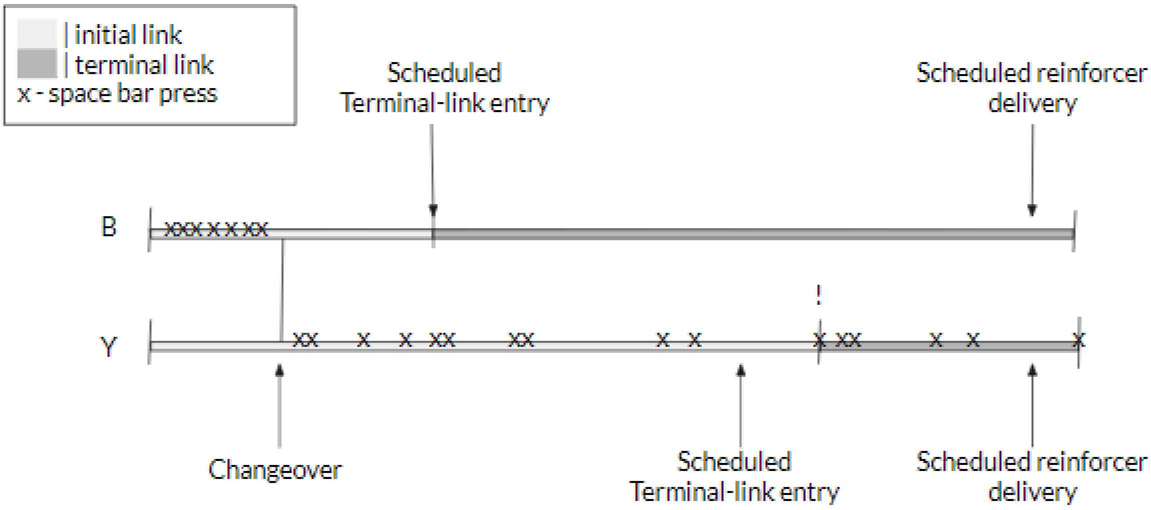


Figure 6. Niu

Best IC Values