Online Appendix For *Complexity and Procedural Choice*

James Banovetz* Ryan Oprea†

A Theoretical Benchmarks

In this Appendix, we derive two benchmarks of interest for our experiment. In Appendix A.1, we derive the critical value for exploration in our bandit problem. In Appendix A.2 we derive the optimal random transition probability for the 2-state partially randomized procedure in Section 2. These derivations closely follow Börgers & Morales (2004) and we refer the reader there for further analysis.

A.1 Critical value for exploration

We can find a critical value such that a decision maker would never explore. Intuitively, if x is close to 1, the cost of exploration (the risk of earning a zero) outweighs benefit (the chance of earning one). With discount rate δ, the condition is:

$$\left(\frac{x}{1-\delta}\right) \geq \frac{1}{3}\left(\frac{\delta x}{1-\delta} + \frac{x}{1-\delta} + \frac{1}{1-\delta}\right) \tag{1}$$

Setting these two equal, we can solve for the critical value.

$$\overline{x} = \frac{1}{1+(1-\delta)} \tag{2}$$

For values larger than $\overline{x}$, it is optimal to remain on the initial arm and never explore.

A.2 Optimal Transition Probability for Partially Random 2-State Procedure

Consider first the value function when the arms are $(x, 0)$ and the DM's strategy dictates a pull of the initial arm:

$$V_0 = x + \delta\big[(1-p)V_0 + p\delta V_0\big] \tag{3}$$

*Banovetz: The Brattle Group, One Beacon Street, Suite 2600, Boston MA 02108, https://www.orbitz.com/trips james.banovetz@gmail.com;

†Oprea (corresponding author): Economics Department, University of California, Santa Barbara, Santa Barbara, CA, 93106, roprea@gmail.com.

The DM's first choice yields a payoff of x, and we discount future payoffs by δ. In the next period, with probability $(1-p)$ she pulls the initial arm again (leaving her state unchanged). With probability p she pulls the other arm, which pays zero; she then returns to the initial arm. Solving for V_0:

$$V_0 = \frac{x}{(1-\delta)(1+\delta p)} \tag{4}$$

Next, consider the value function when the arms are (x, x):

$$V_x = \frac{x}{1-\delta} \tag{5}$$

If both arms pay x, then the DM's pattern of choices do not affect payoffs. Finally, consider the value function when the arms are $(x, 1)$ and the DM's strategy dictates a pull of the initial arm:

$$V_1 = x + \delta \left[(1-p)V_1 + p\left(\frac{1}{1-\delta}\right) \right] \tag{6}$$

The DM pulls the initial arm, earning x. She then stays on the initial arm with probability $(1-p)$; with probability p, she pulls the other arm and stays forever. Solving for V_1:

$$V_1 = \frac{(1-\delta)x + \delta p}{(1-\delta)(1-\delta+\delta p)} \tag{7}$$

In the optimal 2-State strategy, when the DM plays the initial arm, she cannot remember if she has tried the other arm (or what it pays). When her strategy dictates a pull of the initial arm, her value function is the average of V_0, V_x, and V_1.

$$V = \frac{1}{3}\left[\frac{x}{(1-\delta)(1+\delta p)} + \frac{x}{1-\delta} + \frac{(1-\delta)x + \delta p}{(1-\delta)(1-\delta+\delta p)} \right] \tag{8}$$

We can maximize this function with respect to p to find the optimal experimentation probability. Note that maximizing V with respect to p is the same as maximizing:

$$M = \frac{x}{1+\delta p} + \frac{(1-\delta)x + \delta p}{(1-\delta+\delta p)} \tag{9}$$

To find the maximum, we consider the first and second derivatives.

$$\frac{dM}{dp} = -\frac{\delta x}{(1+\delta p)^2} + \frac{\delta(1-\delta)(1-x)}{(1-\delta+\delta p)^2} \tag{10}$$

$$\frac{d^2M}{dp^2} = \frac{2\delta^2 x}{(1+\delta p)^3} - \frac{2\delta^2(1-\delta)(1-x)}{(1-\delta+\delta p)^3} \tag{11}$$

The second derivative is negative when:

$$p < \frac{(1-\delta)^{\frac{1}{3}}(1-x)^{\frac{1}{3}} - (1-\delta)x^{\frac{1}{3}}}{\delta(x^{\frac{1}{3}} - (1-\delta)^{\frac{1}{3}}(1-x)^{\frac{1}{3}})} \tag{12}$$

Recall that p must be greater than zero. Consider the numerator and the denominator separately. The numerator is positive when:

$$(1-\delta)x^{\frac{1}{3}} < (1-\delta)^{\frac{1}{3}}(1-x)^{\frac{1}{3}} \tag{13}$$

$$x < \frac{1}{1+(1-\delta)^2} \tag{14}$$

Note that his value is strictly greater than the critical value in (2). As a result, the numerator will be positive for all values of x such that an DM would consider exploration. The denominator is positive when:

$$x^{\frac{1}{3}} > (1-\delta)^{\frac{1}{3}}(1-x)^{\frac{1}{3}} \tag{15}$$

We can set these equal to solve for a second critical value.

$$\underline{x} = \frac{1-\delta}{2-\delta} \tag{16}$$

For values $x \geq \overline{x}$, we know that $p = 0$. For values $\underline{x} < x < \overline{x}$, we can solve for the optimal p by setting the first derivative in 10 equal to zero and solving for p:

$$p = \frac{\sqrt{1-\delta}\sqrt{1-x} - (1-\delta)\sqrt{x}}{\delta\sqrt{x} - \delta\sqrt{1-\delta}\sqrt{1-x}} \tag{17}$$

What about values of x less than the critical value in (16)? Consider a case where $p = 1$, but the first derivative is still positive (i.e., the maximum of M is at the upper boundary of p). In this situation, the following inequality holds:

$$\frac{x}{(1+\delta)^2} < \frac{\delta(1-\delta)(1-x)}{(1-\delta+\delta)^2} \tag{18}$$

$$x < \frac{1-\delta}{\frac{1}{(1+\delta)^2} + 1 - \delta} \tag{19}$$

The value in (19) is strictly greater than $\underline{x}$. Thus, for all values of x less than the critical value in (16), the optimal experimentation probability is $p = 1$. Note, however, that (19) also established that there is a range of x such that $\underline{x} < x < \overline{x}$ and $p = 1$, i.e., p as define in (17) may be greater than 1. Thus, the optimal experimentation probability must be defined piece-wise:

$$p = \begin{cases} 0 & \text{if } x \geq \overline{x} \\ \min\left\{\frac{\sqrt{1-\delta}\sqrt{1-x}-(1-\delta)\sqrt{x}}{\delta\sqrt{x}-\delta\sqrt{1-\delta}\sqrt{1-x}}, 1\right\} & \text{if } \underline{x} < x < \overline{x} \\ 1 & \text{if } x \leq \underline{x} \end{cases} \tag{20}$$

B Diagnostic Treatments

The experiment includes two diagnostic treatments, designed to better understand the mechanism driving our primary ST/NST treatment comparison.

State Tracking + Distraction (ST-D): In this treatment, subjects are assigned the ST treatment, but between each period the software flashes a random letter on subjects' screens. At the end of each five period block, subjects are asked to type these five letters as part of their payment. In addition to per period payoffs of 0, 65 or 100 points in the bandit task, subjects receive 300 payment points every time they correctly type the five-letter sequence. Instead of being paid for each point in excess of 700 (as in the other treatments), they are paid for each point in excess of 1000.[1] The treatment was run using 60 Prolific subjects in April and May 2020.

Estimated switching probabilities from the second half of the sessio are plotted in graphs analogous to Figure **??** in the top half of Figure 1. There are two main findings from this treatment. First is that subjects' procedures are considerably more likely to be impossible to identify in ST+D than in ST (23% rather than 13%), suggesting a higher rate of noisy mistakes due to the increased task difficulty of the treatment. Second is that conditional on being typed, procedural choice is virtually identical in ST-D and ST. Subjects choose 1/2/3/4 state rules 44%/15%/27%/14% of the time in ST and 48%/13%/26%/13% of the time in ST-D. Unsurprisingly we cannot reject the hypothesis that these distributions are identical via a Wilcoxon test ($p = 0.8$). The results thus suggest that changes to task difficult lead to noisier behavior rather than systematic changes in the procedures subjects use.

No State Tracking - State Tracking (NST-ST): In this treatment subjects are given the exact same instructions used in the NST treatment and are assigned that treatment for the first ten (of twenty total) tasks. After task ten, subjects are given new instructions informing them that they will no longer be shown question marks under arms they haven't yet selected or payoff amounts under arms they have. That is, they are told they will play the rest of the tasks under the ST treatment. This change of treatment is a surprise to subjects (it is not discussed in the earlier instructions). Payments and incentives are exactly as in the NST and ST treatments. The treatment was run using 61 Prolific subjects in January 2022. Estimated switching probabilities from the second half of the session are plotted in graphs analogous to Figure **??** in the top half of Figure 1. The main findings from this treatment are discussed in detail in Section 4.5.

C Instructions to Subjects

In this Appendix we reproduce the instructions to subjects. These were deployed in HTML in the experiment and unfolded progressively interspersed with comprehension quiz questions. In Appendix C.1 we reproduce instructions from our ST treatment and in Appendix C.2 instructions from our NST treatment.

C.1 State Tracking Treatment

1. **Introduction**

    - We will start by providing you with **INSTRUCTIONS** for the study.

[1]Also, because these sessions are longer, subjects are paid a larger base pay of \$5. rather than the \$2.50 from other treatments.

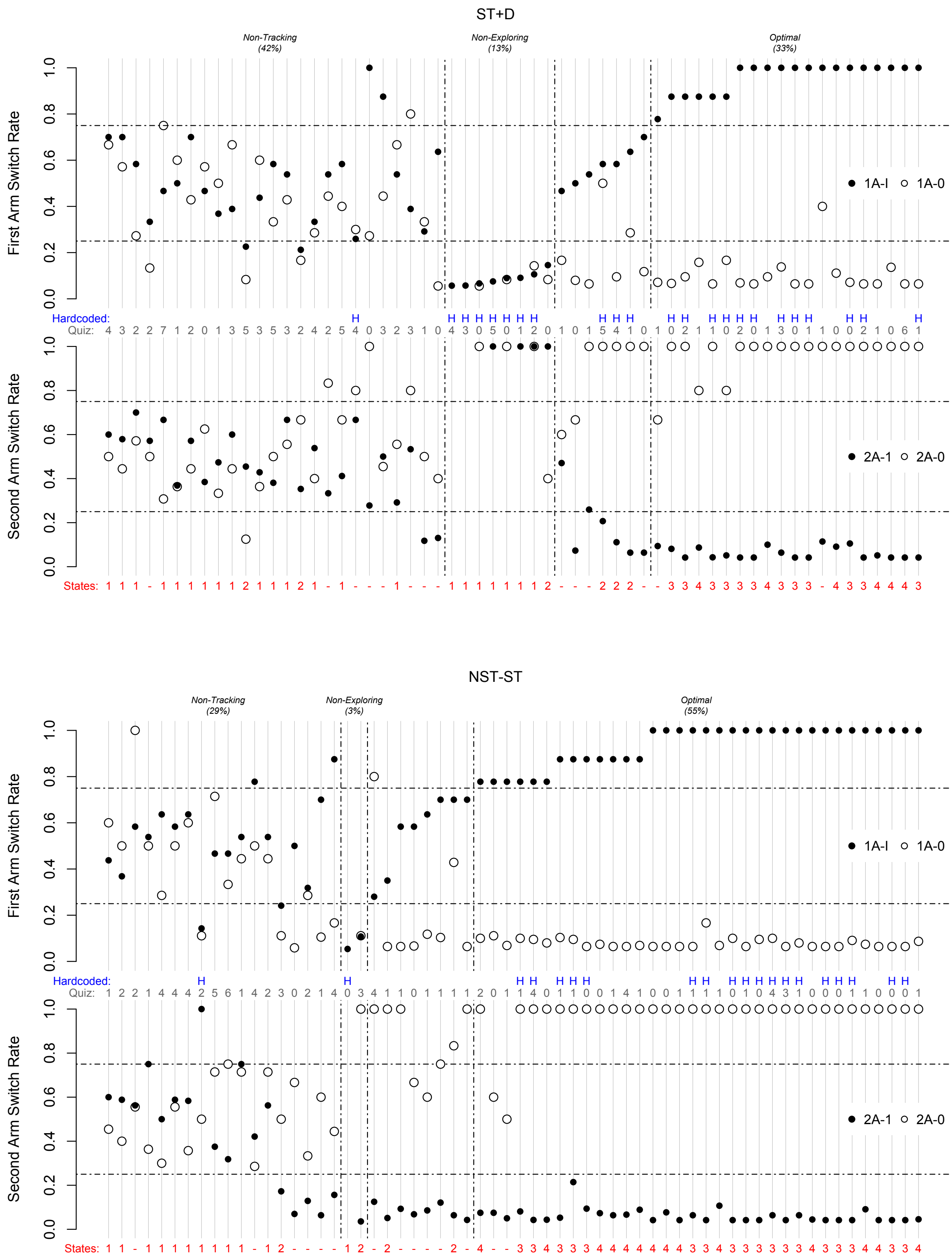
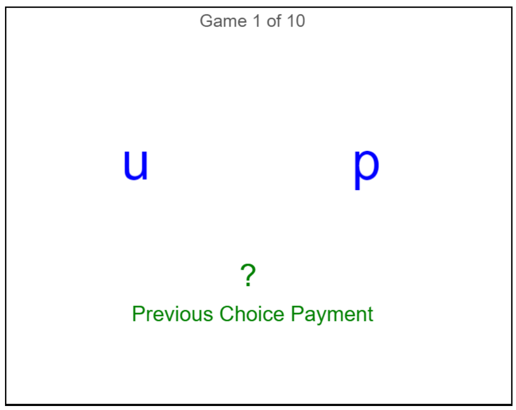


Figure 1: Switching probabilities for each subject in our diagnostic treatments, ST+D (upper panel) and NST-CL (lower panel). *Notes: The upper plot of each panel shows switching behavior after play of the initial arm, the lower plot after play of the second arm. Numbers at the bottom of each panel shows the estimated number of states in the procedure the subject used. Numbers between plots in each panel show the number of of errors the subject made in a comprehension quiz prior to the experiment, and the letter 'H' between plots in each panel designates a subject who made consistent ("hardcoded") initial choices.*

- We will ask you **QUESTIONS** to check that you understand the instructions. You should be able to answer all of these questions correctly.
- Please read and follow the instructions closely and carefully.
- If you **COMPLETE** the main parts of the study, you will receive a **GUARANTEED PAYMENT** of **$2.50**.
- In addition, your **CHOICES** in the GAME portion of the study will result in **PERFORMANCE-BASED EARNINGS**. You will play in **TWENTY (20) GAMES** worth **REAL MONEY**. Your **AVERAGE** points from **ALL TWENTY GAMES** will be converted into an additional payment.
- After you finish the instructions, you will have a chance to play several **PRACTICE GAMES** before you play for real money.

2. **Two Options to Choose Between**



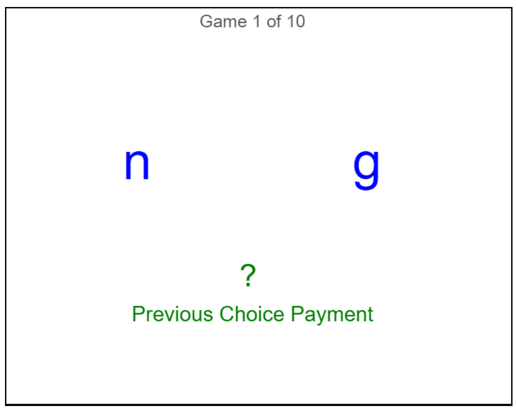
- The experiment is divided into twenty **GAMES**, each of which is is divided into several **CHOICES**.
- In each game, you will repeatedly choose between **TWO OPTIONS** that we will call **EARLIER-LETTER** and **LATER-LETTER**, represented by two letters on your screen.
  - In the example above, the Earlier-Letter option is represented by '**p**' and the later-letter by '**u**' (because '**p**' comes **earlier** than '**u**' in the alphabet).
- Your **FIRST** choice in a game will **ALWAYS PAY 65** points. This is true whether you choose Earlier-Letter or Later-Letter first.
  - For example, suppose your **first** choice were **Earlier-Letter**. Then you would know that Earlier-Letter would pay **65 points every time** you choose it for the rest of the game.
- After your first choice, the **OTHER OPTION** will pay a **VALUE** of either **0**, **65**, or **100**. This value is **INDEPENDENTLY** and **RANDOMLY** determined by the computer

before the game begins. Each value (65, or 100) is **EQUALLY LIKELY** to be selected. It remains the **SAME WITHIN A GAME**.

  - For example, suppose your **first** choice were **Later-Letter**. Then **Earlier-Letter** would be equally likely to pay (0,65, or 100). If you choose Earlier-Letter and it pays 100, then you would know that Earlier-letter would pay **100 points every time** you choose it for the rest of the game.

- However, the value of each option **CHANGES BETWEEN GAMES**: once a game ends, payments reset. Your first choice (either Earlier-Letter or Later-Latter) will pay 65, and the other option will get a new random value.

3. **Typing Letters to Make Choices**



- For each **CHOICE**, each of the **TWO OPTIONS** will require you to **TYPE A LETTER**.

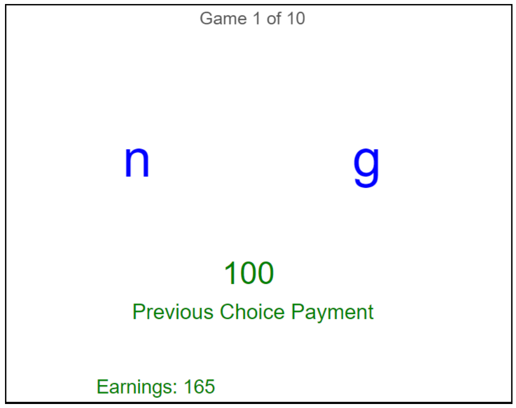
  - In the example above, you would need to type '**n**' (lower case 'N') or '**g**' (lower case 'G') to make a choice.

- These **LETTERS** will **RANDOMLY CHANGE** ('a' to 'z') for each choice and be shown in a **RANDOM ORDER** (left or right) on your screen.

- The **EARLIER LETTER** in the alphabet will always represent the Earlier-Letter option; the **LATER LETTER** in the alphabet will always represent the Later-Letter option.

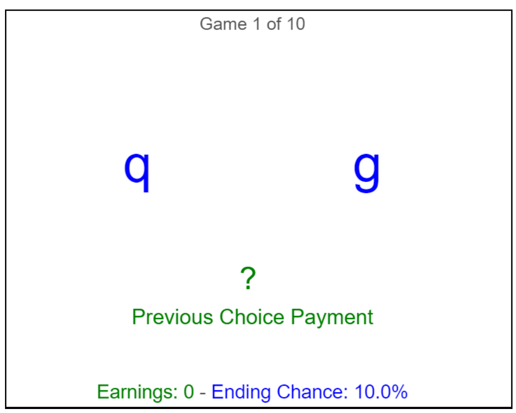
  - In the example above, typing '**g**' will select the Earlier-Letter option (and give you the Earlier-Letter payment) while typing '**n**' will select the Later-Letter option (and give you the Later-Letter payment).

4. **Tracking Payments**



- After you choose an option, the **PAYMENT** you earned (**0**, **65**, or **100**) for that choice appears in the middle of the screen in green. This value always represents your **PREVIOUS CHOICE'S** payment.
  - In the example above, the previous choice paid **100** points.
- Your **EARNINGS** are cumulative for **ALL YOUR CHOICES** so far in the game and appear at the bottom in green.
  - This number is the **sum** of all your payments so far in the game.
  - In the example above, the choices have paid a total of **165** points so far.

5. **Blocks of Choices**



- The **NUMBER OF CHOICES** you're allowed to make in any game is **RANDOM**. Every time you make a choice, there is a **10% CHANCE** that the computer will make it the **LAST** (paying) choice of the game.
  - A 10% chance that the game ends each choice means that on **average** there will be **10 choices** in the game.
  - Many games will be **shorter**, but others will be **much longer**.
  - The probability each choice is the last **does not depend** on how many choices you have already made. Every choice is equally likely to be the last one that counts.

- In each game, you will always make your choices in **BLOCKS OF FIVE**.
- After every block of five the computer will tell you whether the game actually **RANDOMLY ENDED** during that block. If the computer randomly ended the game during the block (before the last choice of the block), any choices you made **AFTER THE LAST** choice in the block **WON'T COUNT** for payment.
  - Example: If you make **five choices** in a block and the computer randomly **ended** the game on the **third choice** of the block, choices **1, 2 and 3 in the block will count** for payment and choices **4 and 5 in the block will not count** for payment.
- When you have made the **FINAL CHOICE** in a game, the computer will inform you that this has happened and you will start a **NEW GAME**. When a new game starts, the **VALUES WILL CHANGE** for the Earlier-Letter and Later-Letter options. There is no connection between games – each game will be brand new.

6. **Other Instructions**
   - You will play in two practice games to familiarize yourself with the software before you play games for real money.
   - Because part of the experiment tests your memory, please do not use external tools (e.g., pencil and paper) to assist during the experiment.
   - Please do not unnecessarily refresh your browser, as doing so can make the software unstable. Qualtrics tracks your refreshes–excessive refreshes will void your bonus payment.

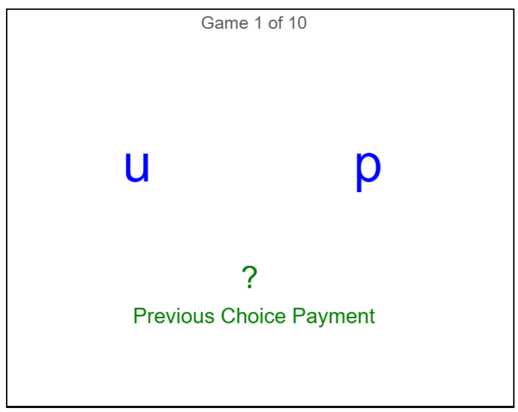
7. **Cash Payments**
   - You will be paid $2.50 for finishing the experiment. If you decide to leave before finishing, you will forfeit this amount.
   - In addition, you will potentially earn a **PERFORMANCE-BASED BONUS**.
   - Your **POINTS** from **ALL TWENTY (20) GAMES** will be **AVERAGED**.
   - If your **AVERAGE** point total is **GREATER THAN** 700, you will earn a **BONUS**.
     - For every point you earn (on average) greater than 700, you will be paid $0.03 (three cents)
     - For example, if you average 950 points, your bonus would be: (950 - **700**) * $0.03 = $7.50
     - For example, if you average 600 points, you would not earn a bonus.

C.2 No State Tracking Treatment

1. **Introduction**
   - We will start by providing you with **INSTRUCTIONS** for the study.

- We will ask you **QUESTIONS** to check that you understand the instructions. You should be able to answer all of these questions correctly.
- Please read and follow the instructions closely and carefully.
- If you **COMPLETE** the main parts of the study, you will receive a **GUARANTEED PAYMENT** of **$2.50**.
- In addition, your **CHOICES** in the GAME portion of the study will result in **PERFORMANCE-BASED EARNINGS**. You will play in **TWENTY (20) GAMES** worth **REAL MONEY**. Your **AVERAGE** points from **ALL TWENTY GAMES** will be converted into an additional payment.
- After you finish the instructions, you will have a chance to play several **PRACTICE GAMES** before you play for real money.

2. **Two Options to Choose Between**



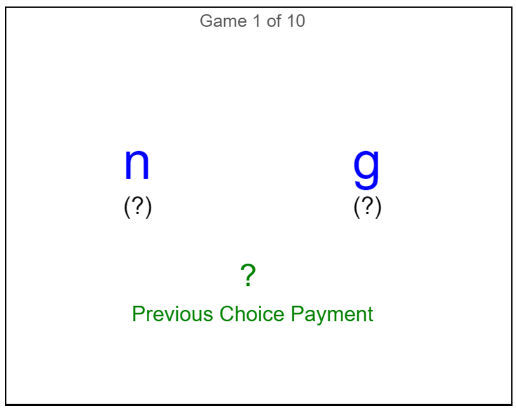
- The experiment is divided into twenty **GAMES**, each of which is is divided into several **CHOICES**.
- In each game, you will repeatedly choose between **TWO OPTIONS** that we will call **EARLIER-LETTER** and **LATER-LETTER**, represented by two letters on your screen.
  - In the example above, the Earlier-Letter option is represented by '**p**' and the later-letter by '**u**' (because '**p**' comes **earlier** than '**u**' in the alphabet).
- Your **FIRST** choice in a game will **ALWAYS PAY 65** points. This is true whether you choose Earlier-Letter or Later-Letter first.
  - For example, suppose your **first** choice were **Earlier-Letter**. Then you would know that Earlier-Letter would pay **65 points every time** you choose it for the rest of the game.
- After your first choice, the **OTHER OPTION** will pay a **VALUE** of either **0**, **65**, or **100**. This value is **INDEPENDENTLY** and **RANDOMLY** determined by the computer

before the game begins. Each value (65, or 100) is **EQUALLY LIKELY** to be selected. It remains the **SAME WITHIN A GAME**.

  - For example, suppose your **first** choice were **Later-Letter**. Then **Earlier-Letter** would be equally likely to pay (0,65, or 100). If you choose Earlier-Letter and it pays 100, then you would know that Earlier-letter would pay **100 points every time** you choose it for the rest of the game.

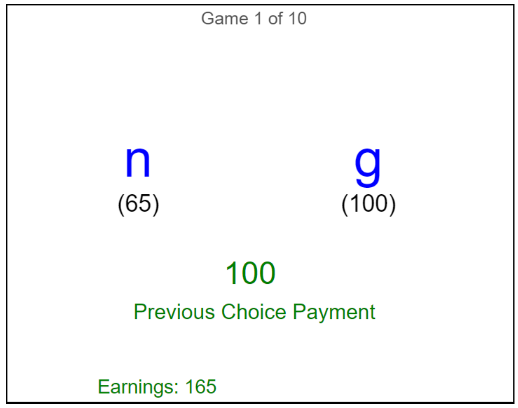
- However, the value of each option **CHANGES BETWEEN GAMES**: once a game ends, payments reset. Your first choice (either Earlier-Letter or Later-Latter) will pay 65, and the other option will get a new random value.

3. **Typing Letters to Make Choices**



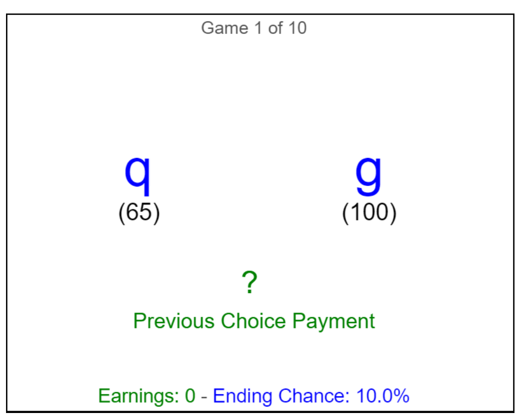
- For each **CHOICE**, each of the **TWO OPTIONS** will require you to **TYPE A LETTER**.
  - In the example above, you would need to type '**n**' (lower case 'N') or '**g**' (lower case 'G') to make a choice.
- These **LETTERS** will **RANDOMLY CHANGE** ('a' to 'z') for each choice and be shown in a **RANDOM ORDER** (left or right) on your screen.
- The **EARLIER LETTER** in the alphabet will always represent the Earlier-Letter option; the **LATER LETTER** in the alphabet will always represent the Later-Letter option.
  - In the example above, typing '**g**' will select the Earlier-Letter option (and give you the Earlier-Letter payment) while typing '**n**' will select the Later-Letter option (and give you the Later-Letter payment).
- The **QUESTION MARKS** (in parentheses) under each letter indicates that you have not yet tried that option.
  - In the example above, the '**(?)**' under '**g**' indicates Earlier-Letter has not been tried
  - In the example above, the '**(?)**' under '**n**' indicates Later-Letter has not been tried.

4. **Tracking Payments**



- After you choose an option, the **PAYMENT** you earned (**0**, **65**, or **100**) for that choice appears in the middle of the screen in green. This value always represents your **PREVIOUS CHOICE'S** payment.
  - In the example above, the previous choice paid **100** points.
- Your **EARNINGS** are cumulative for **ALL YOUR CHOICES** so far in the game and appear at the bottom in green.
  - This number is the **sum** of all your payments so far in the game.
  - In the example above, the choices have paid a total of **165** points so far.
- After you have **CHOSEN** an option, the **AMOUNT** it pays appears **BELOW**. This number will remain until the game ends and a new one begins
  - In the example above, Earlier-Letter was tried and paid **65**
  - In the example above, Later-Letter was tried and paid **100**

5. **Blocks of Choices**



- The **NUMBER OF CHOICES** you're allowed to make in any game is **RANDOM**. Every time you make a choice, there is a **10% CHANCE** that the computer will make it the **LAST** (paying) choice of the game.

  - A 10% chance that the game ends each choice means that on **average** there will be **10 choices** in the game.
  - Many games will be **shorter**, but others will be **much longer**.
  - The probability each choice is the last **does not depend** on how many choices you have already made. Every choice is equally likely to be the last one that counts.
- In each game, you will always make your choices in **BLOCKS OF FIVE**.
- After every block of five the computer will tell you whether the game actually **RANDOMLY ENDED** during that block. If the computer randomly ended the game during the block (before the last choice of the block), any choices you made **AFTER THE LAST** choice in the block **WON'T COUNT** for payment.
  - Example: If you make **five choices** in a block and the computer randomly **ended** the game on the **third choice** of the block, choices **1, 2 and 3 in the block will count** for payment and choices **4 and 5 in the block will not count** for payment.
- When you have made the **FINAL CHOICE** in a game, the computer will inform you that this has happened and you will start a **NEW GAME**. When a new game starts, the **VALUES WILL CHANGE** for the Earlier-Letter and Later-Letter options. There is no connection between games – each game will be brand new.

6. **Other Instructions**

- You will play in two practice games to familiarize yourself with the software before you play games for real money.
- Please do not unnecessarily refresh your browser, as doing so can make the software unstable. Qualtrics tracks your refreshes–excessive refreshes will void your bonus payment.

7. **Cash Payments**

- You will be paid $2.50 for finishing the experiment. If you decide to leave before finishing, you will forfeit this amount.
- In addition, you will potentially earn a **PERFORMANCE-BASED BONUS**.
- Your **POINTS** from **ALL TWENTY (20) GAMES** will be **AVERAGED**.
- If your **AVERAGE** point total is **GREATER THAN 700**, you will earn a **BONUS**.
  - For every point you earn (on average) greater than 700, you will be paid $0.03 (three cents)
  - For example, if you average **950** points, your bonus would be: (**950** - **700**) * $0.03 = $7.50
  - For example, if you average **600** points, you would not earn a bonus.

References

Börgers, T. & Morales, A. (2004), 'Complexity Constraints in Two-Armed Bandit Problems: An Example'.