Revisit the Epiphany Learning Behavior in the Two-person Beauty Contest Game

Wei James Chen[a,*]

[a]*Department of Business Administration, National Central University, No.300, Zhongda Rd., Zhongli Dist., Taoyuan City 32001, Taiwan (R.O.C.)*

Abstract

People often fail to make optimal choices initially, but learn to act optimally in the long run. Thus, investigating learning behaviors is key to understanding human behavior. In economics and psychology, reinforcement learning (RL) have received much attention in the literature. However, not all behavior appears to conform to the gradual behavioral convergence that is a central feature of RL. In some cases, learning appears to happen all at once. Previous studies have documented this phenomenon in various experimental setting, and call it epiphany learning (EL). Nevertheless,, in these studies, subjects were restricted to only use pure strategy, causing potential biased estimation of the proportion of epiphany learners. In this paper, we investigate the EL using a mixed strategy method and apply a computational model of EL we previously proposed in a two-person beauty contest experiment. We identify 45% of our subjects as being epiphany learners with the mixed strategy method, which is indeed higher than previous estimates (ranging from 35 to 39%). This result indicates that the pure strategy method is potentially underestimating the proportion of epiphany learners.

Keywords: Two-person beauty contest, Weak dominance, Lab experiment, Eureka learning, Bounded rationality
JEL: C91, D83

1. Introduction

Over the last few decades, learning has been a central question in economics. It has proven to be valuable in explaining both empirical and experimental data. Economists study learning to understand how people learn to act optimally in the long run, since they often fail to do so initially. As Camerer and Ho (1999) noted, economists study learning to answer the following question: "how does an equilibrium arise in a noncooperative game?" The majority of researchers believed that the answer to this question is reinforcement learning

* Corresponding author. *Email address:* `thisis.chenwei@gmail.com` (Wei James Chen) *URL:* `http://u.osu.edu/chen.2731/` (Wei James Chen)

(RL). Cross (1973) and Arthur (1991) first adopted the idea of RL from psychology to study decision-making. Later Mookherjee and Sopher (1994) studied RL in a noncooperative game setting. In the subsequent years, many studies focused on bringing psychological effects such as spillover (Roth and Erev, 1995), aspiration level (Börgers and Sarin, 2000), small sampling and weighting (Arifovic et al., 2006), inertia (Biele et al., 2009), bounded memory (Chen et al., 2011) and observation on another person's action and/or payoff (Hanaki et al., 2018) into RL to better accommodate experimental data.

All the modifications above share a fundamental feature of the RL model, namely the predicted choice probabilities gradually change over time. However, in many cases, the data do not have this feature. In psychology, this has been seen in insight problems such as the nine-dot problem, and the triangle problem (Metcalfe, 1986; Metcalfe and Wiebe, 1987). Subjects' learning behavior in these problems feature a strong discontinuity in the feeling of knowing (knowing the solution, see Pols (2002) for a summary), and the choice probabilities sequence has a discontinuous jump. This discontinuity has been linked to a sudden increase in heart rate (Jausovec and Bakracevic, 1995) and pupil dilation (Nassar et al., 2012).

Moreover, in experimental economics, two previous studies (Dufwenberg et al., 2010; McKinney and Van Huyck, 2013) also found such learning phenomenon in two extensive form games and named it epiphany learning (EL) or eureka learning. And our previous work (Chen and Krajbich, 2017) found EL in a decision-making problem. In these studies, the existence of an epiphany moment were established by the discontinuity in the choice sequence (learning curve) or by model comparison. For example, McKinney and Van Huyck (2013) applied the Pettitt change-point analysis to test if there is a structural change-point in subject's choice sequence, and Chen and Krajbich (2017) showed that a computational model of EL can better fit a significant proportion of subject's choice sequence than a standard RL model.

However, the choice sequences used to identify epiphany in the previous literature were all sequences of discrete choice, which may lead to a biased identification. To see why, suppose a decision-maker is actually playing a mixed strategy in a series of choice between L and R, but we were only able to observe a sequence of discrete choice. Since the realization of her mixed strategy sequence is random, a choice sequence that has a discontinuity (EL) might come from a gradually increasing mixed strategy sequence (RL), and a choice sequence that has no discontinuity (RL) might come from an abruptly increasing mixed strategy sequence (EL). In other words, the estimated EL might be RL and vice versa. Therefore, in this study, we allow subjects to use mixed strategy so that we can avoid the above issue of identification.

We study EL using the 2-person beauty contest (2BC) game (Grosskopf and Nagel, 2008), the same game we used in the previous study (Chen and Krajbich, 2017). The 2BC is a special case of the p-beauty contest (Nagel, 1995), with only 2 players. Both players guess an integer between 0 and 100. The winner is the player whose guessed number is closet to p multiplied by the average guess (the target number), where $0 < p < 1$. Unlike in the usual p-beauty contest with more than two players, 0 is a weakly dominant strategy in the 2BC, not just a Nash equilibrium strategy. Past research (Grosskopf and Nagel, 2008, 2009; Chen and Krajbich, 2017) has shown that most people fail to realize that 0 is the optimal strategy initially, but eventually, many figure it out. We applied a previously proposed EL model (Chen and Krajbich, 2017) to test the hypothesis that this learning process is an epiphany type of learning.

The key feature of this EL model (please refer to the Appendix A for details of the EL and RL model) is a dynamic evidence (in favor of 0) variable, $ev(t)$, with $ev(0) = 0$. Here the t is a time variable for rounds. At the end of each round, a decision-maker either aggregates positive evidence d ($ev(t) = ev(t-1) + d$), or negative evidence $-d$ ($ev(t) = ev(t-1) - d$), depending on the payoff history she observes. Once the $ev(t)$ hits a predefined threshold (1 or -1), she gains a certain type of epiphany, and believe that the strategy corresponding to that epiphany is optimal. Before she has the epiphany, she chooses the strategy that she initially believes to be optimal with probability q_1, and after the epiphany, she chooses the strategy that she now believes to be optimal with probability q_2 (here d, q_1 and q_2 are parameters in the EL model). This model aims to capture the epiphany component in the learning behavior. An epiphany will occur only if enough evidence has been accumulated, and unless an epiphany has occurred, the decision-maker will not change her strategy at all. Moreover, once the epiphany occurs, the decision-maker will not change her behavior again unless a different kind of epiphany occurs.

To preview our results, we found that subjects' choices were mostly (66%) mixed strategy[1]. Besides, 42 out of 62 subjects learned the optimal strategy during the experiment. Among these 42 subjects, we found 28 of them better fitted by the EL model, compared to the RL model. Moreover, subjects' choice prediction errors[2] in early trials can positively predict the chance of them being epiphany learners later on. This result suggests that EL seems to be triggered by bigger mistakes (higher prediction errors) in the learning phase.

The remainder of this paper is organized as follows: Section 2 summarizes previous literature. Section 3 introduces the 2BC game. The experimental design is in Section 4, and Section 5 demonstrates our experimental results. Lastly, Section 6 discusses our results and concludes.

2. Previous Literature

2.1. Epiphany

In psychology, Wallas (1926) first defines the concept of insight learning as the combination of incubation and illumination (insight). In the incubation stage, the problem solver does not try to solve the problem or at least not do so consciously. Researchers often assume the existence of incubation or create a period of time for subjects to incubate. Dodds et al. (2003) summarize these experiments and find that, first, subject's performance is positively correlated with the length of incubation. Second, providing useful clue improves the quality of incubation. Both results suggest the usefulness of incubation.

On the other hand, there are several different definitions of insight. Pols (2002) summarizes three main elements of this concept. First, it is a transition that has a major impact on

[1] We define a choice as mixed strategy if the maximum probability put on a single number is less than 2/3.
[2] Defined as subject's decision (realized from the mixed strategy) minus the target number in the same trial.

the problem solver's conception of that problem. Second, insight is sudden. Third, the new understanding is more appropriate than the previous understanding. Therefore, empirically, researchers often ask subjects to report their feeling of knowing (the solution) during the experiment, and insight is often identified by a strong discontinuity in the feeling of knowing (Pols, 2002; Metcalfe and Wiebe, 1987). Jausovec and Bakracevic (1995) even find a sudden increase of heart rate right before an insight was reached.

Hélie and Sun (2010) combine incubation and insight in a theoretical model. In their theory, insight is reached when the internal confidence level reaches a threshold, similar to the EL model we used in this study. However, this theory is designed for insight problem (such as the nine-dot problem) solving, not for the decision-making problem, so it does not rely on explicit feedback from the environment.

There is another branch of study called all-or-none (AON) learning in psychology that is related to our work. Bower (1961) proposes a model with two states of mind (conditioned to the correct response or not), which can be understood as not having an epiphany and having an epiphany in our context. A Markov chain is used to describe the transition between the two states. Specifically, the probability of getting the epiphany is a fixed probability c, and once one has the epiphany, the probability of staying in that state of mind is 1. Murre (2014) examines the AON learning in a foreign vocabulary learning experiment. He shows that the individual learning curve can be well fitted by an exponential (S-shaped) function. Note that the AON learning model and the EL model we used both produce S-shaped learning curves, but the timing of insight is random in the AON learning model.

In economics, Dufwenberg et al. (2010) use the game of 21 to test the existence of EL[3]. This game is a two-person zero-sum sequential game of perfect information, so the optimal strategy can be solved by backward induction. Half of the subjects first play the game of 6 for 5 rounds, which is a similar but simpler version (with the identical optimal strategy) of the game of 21, and then play the game of 21 for another 5 rounds. The other half of the subjects play the two games in the opposite order. The authors show that most subjects fail to play the optimal strategy in the game of 21, unless they first play the game of 6. Moreover, the authors demonstrate that many subjects do not gradually learn the optimal strategy in the game of 21, but reach that strategy all of a sudden. They identify 36% of their subjects as epiphany learners.

McKinney and Van Huyck (2013) test the existence of eureka (epiphany) learning in the lab using the Nim game. The Nim game is also a two-person zero-sum sequential game of perfect information, so the subgame perfect equilibrium strategy is the optimal strategy. Subjects play a sequence of 92 Nim games with different difficulties against a computer algorithm that always play the optimal strategy. The authors find that more than 35% of their subjects can be classified as eureka learners, by using the Pettitt change-point analysis to find a statistically significant change-point (epiphany moment) in the learning curve.

Our paper differs from the above two papers in three points. First, the 2BC game is not a extensive-form game. In a extensive-form game, subjects may need to first gain the

[3] In the game of 21, two players take turns adding one or two to the current number and start with 0 as the initial number. Whoever reaches the number 21 wins.

epiphany of using the backward induction, and then realize that there is an optimal strategy from a second epiphany (Dufwenberg et al., 2010). This problem complicates the definition of epiphany. Second, subjects were allowed to report mixed strategy. Therefore, we can uncover subjects' beliefs about what strategy is optimal, and if subjects play the optimal pure strategy, we can be more confident that they are playing the optimal strategy, and are not just choosing that action by chance. Third, in the above two papers, subjects either play against another subject whose strategy is changing over time or an optimal algorithm. In the former case, subjects may notice that their opponent's strategy is converging to the optimal strategy, and in the latter case, subjects can simply observe and learn the optimal strategy from their opponent's previous move. In both cases, subjects may act as if they realize the optimal strategy, without knowing it. In our design, subjects play against a random selection from a database of subjects who previously played the same game, so their opponent's strategy is neither converging nor always optimal[4]. Besides, by the common definition of epiphany, the EL process should mainly be internal. Our design makes sure the learning process is less affected by the external environment.

Lastly, in our previous paper (Chen and Krajbich, 2017) we used the same one-shot game (the 2BC game) as in this study to test the existence of EL. Also, subjects were playing against a fixed mixed strategy, constructed from a database of previous experiment, so learning from their opponent's strategy was limited. However, subjects were still only allowed to use pure strategy, so using choice data along to identify EL were potentially biased. There, we addressed this issue by giving subjects an option to "commit" to an action they chose at any round, namely locked into that action for the rest of the experiment. Subjects' commitment was designed as a proxy to their epiphany moment, so we would be able to correctly identify who had the epiphany. We found 25 out of 59 subjects who committed to the optimal strategy, and the vast majority of them (23 out of 25) were identified as epiphany learners. However, this commitment method raised another issue, we found another group of subjects (23 out of 59) who committed to other non-optimal strategies, and because they were not allowed to take back their commitment, we were not able to know if they would eventually have a correct epiphany. In other words, the proportion of epiphany learners we estimated there was a lower bound. Therefore, in this study, we used the mixed strategy method rather than the commitment method to have a more precise estimation of the EL proportion.

2.2. The two-person beauty contest game

Nagel (1995) studies the p-beauty contest in an experimental setting. She finds that different values of p (1/2, 2/3, or 4/3) affect the speed with which subjects' strategies converge to the Nash Equilibrium strategy (playing 0). For example, when $p = 1/2$ (as compared to 2/3), subjects' strategies converge more quickly. She concludes that the depth of reasoning is the key factor for explaining this difference. Some subjects realize that numbers greater than $100 \times p$ are dominated, and others think one step further and realize

[4] By doing so, we do not change the structure of the game. Therefore, this method still preserves incentives for individuals to play optimally.

that if no one chooses those numbers, then numbers greater than $100 \times p^2$ will be dominated as well, and so on. Therefore, the majority of the subjects chose numbers like $100 \times p$, $100 \times p^2$, or $100 \times p^3$, etc (see Ho et al. (1998) for the Level-k model and Camerer et al. (2004) for the Cognitive Hierarchy theory). As a result, we use $p = 0.9$ to make sure that if subjects indeed form their strategy this way, their strategy will converge slowly (with a rate of convergence p) to the dominant strategy in our experiment, distinguishing it from the sudden shift predicted by the EL model.

Grosskopf and Nagel (2008) find that subjects' initial responses to the 2BC game deviate from the weakly dominant strategy (playing 0), even in a professional economist/psychologist subject pool. Grosskopf and Nagel (2009) study this phenomenon in depth. In their experiment, subjects play the 2BC for 10 rounds with either full, partial, or no feedback. They find some evidence of level-k learning, but do not fit their data with the level-k model[5]. In addition, more information in the feedback indeed leads to faster convergence to the dominant strategy. They conclude that subjects learn to adopt the dominant strategy because they observe that their opponents' strategies are converging to 0 over time. We try to limit this channel of learning in our experiment by giving subjects only partial feedback, and letting them play against random draws from a fixed distribution so that the opponent's strategy is not converging.

Chou et al. (2009) study the 2BC to understand why subjects fail to play the dominant strategy in this game. They use an isomorphic transformation of the 2BC[6] and show that subjects are able to choose the dominant strategy in this new game. Therefore, they claim that subjects fail to behave optimally in the 2BC game because they do not understand the game form, which is defined as the relationship between possible choices, outcome, and payoffs. Their results suggest that the only epiphany one needs in the 2BC is that whoever chooses the smaller number wins the game, so the term epiphany is clearly defined in the 2BC.

Nagel et al. (2016) study a "distance-payoff" 2BC, where subjects' payoffs are in proportion to the distance between their choices and the target number. Here, zero is not the optimal choice, so this is not the same game as the usual 2BC. However, they find that subjects' one-shot choice distribution looks like the usual 2BC. Their results suggest that subjects may initially fail to understand the true payoff structure of the 2BC, and thus leads to sub-optimal choices, and the epiphany they need is the realization of the true payoff structure.

3. The Game

The beauty contest game (or p-beauty contest) is a number guessing game. In the usual setting, there are n (≥ 2) players, and each of them submits an integer ranging from 0 to

[5] Gill and Prowse (2016) estimate a structural model of learning based on the level-k reasoning in the general p-beauty contest game and find that it fits better than the experience-weighted attraction (EWA) learning model (Camerer and Ho, 1999), which is a more general case of the RL model. [6] Two subjects are told to choose a number between 0 and 100, and whoever chooses the smaller number wins the game.

100. The average of all players' submissions is computed and then multiplied by a constant p (usually between 0 and 1) to get the target number. Whoever submits a number that is closest to the target number wins the game. In this game, the unique Nash equilibrium strategy is to guess 0, but this is not a dominant strategy when $n > 2$. For example, for $n = 3$ and $p = 0.9$, suppose the 3 guesses are 0, 90, and 100. Then, the target number will be 57, so the closest guess is 90. In fact, this game does not have any (weakly) dominant strategies.

In order to observe the epiphany in learning behavior, we prefer to use a game that has a dominant strategy so that the term epiphany (realizing the dominant strategy) can be clearly defined. Therefore, we employ a special case of the beauty contest game, which is the beauty contest with 2 players, or 2BC. In this game, for any $p \in (0, 1)$, since the average of any two numbers multiplied by p will always be closer to the smaller number, guessing 0 is the unique weakly dominant strategy. Moreover, in order to obtain subjects' detailed belief about what strategy is optimal, we ask subjects to submit a mixed strategy (a distribution over the choice set). Since submitting a mixed strategy over the range between 0 and 100 is too laborious, the range of allowable guesses is reduced to $[0, 10]$.

4. Experimental Design

Ninety-four undergraduate students from the National Taiwan University (NTU) were recruited from the Taiwan Social Science Experimental Laboratory recruiting website. Thirty-two students formed the database (please refer to the Appendix B for this experiment and the screen-shot of the choice screen) and 62 completed the formal experiment (Two sessions were conducted, one with 30 subjects, and the other with 32 subjects).

All sessions of our experiment were programmed in z-Tree (Fischbacher, 2007). Subjects in the formal experiment first played the usual beauty contest ($n > 2$) with $p = 0.9$ for one practice round (no payment), and received feedback about the target number and game result. In this round, all subjects were in a single group, so 0 was not the dominant strategy. Note that we asked subjects to reveal their mixed strategies in this round as well[7]. After the practice round, subjects played the 2BC game ($p = 0.9$) for 20 rounds, and their opponents' choices were drawn from the database. Specifically, for each player, in each round, we drew from the database a random player's choice in a random round between round 1 to 15 [8]. Therefore, subjects faced the exact same game in all 20 rounds. At the end of each round, subjects were told the target number, game result, and payoffs received . In addition to a \$100 NTD (approximately \$ 3.30 USD) show-up fee, subjects received \$10 NTD for each win, nothing for each loss, and in the case of a tie, a digital coin-flip would decide the winner.

[7] The practice round here was intended to make sure subjects understood how to make choices using the experimental interface, but we did not want subjects to learn the optimal strategy in the 2BC in this round.

[8] We did not take all 20 rounds from the database because the percentage of 0 being chosen would appear too often in that case. We did reveal this information to subjects.

5. Results

5.1. Are subjects using mixed strategy?

Since we are using the mixed strategy method to identify epiphany, the first thing we should ask is whether subjects are actually using mixed strategy. Here, we say that a subject is using a mixed strategy if the maximum probability put on a single number is less than 2/3. Overall, subjects were using mixed strategy in 66% of all rounds, excluding rounds that they put all chances on 0. In other words, we did see that subjects were using mixed strategy.

5.2. Are subjects learning?

Next, we sought to see if subjects learned the optimal strategy in the 2BC game. We excluded 20 subjects for the rest of the analysis, because they fail to learn the optimal strategy during the experiment. Eight of them always put at least 95% of the probability weight on 0, indicating that they had already figured out the optimal strategy before the first round, and the rest of them ($n = 12$) never chose the dominant strategy with probability greater than 0.95, suggesting that they never figured out that playing 0 is the weakly dominant strategy[9].

The rest of the subjects showed sign of learning. Wilcoxon rank-sum test rejected the hypothesis that the probability of choosing 0 are equal across the first and the last five rounds (p-value < 0.0001).

5.3. How are they learning?

To understand how subjects learn in the experiment, and distinguish between the EL and the RL model (please refer to the Appendix A for details of the EL and RL model). We need to estimate both models on the individual level because the average learning behavior from EL and RL are similar (Estes, 1956). For example, in Figure 1, we have a group average learning curve from 600 simulated EL subjects with different d values, playing the same game for 20 rounds[10]. However, from this figure, one could easily think that these subjects are a group of reinforcement learners, because the group average learning curve is gradually increasing from 0.1 to 0.7.

5.3.1. Model estimation method

For simplicity, we assume subjects' strategies to be stage-game strategies, and denote their actual choice of a mixed strategy in round t by $s(t)$. Recall that $s(t)$ is a mixed strategy. Thus, it is a distribution over the set $X = \{0, 1, \ldots, 10\}$. Secondly, we assume the estimated parameters in both models to be heterogeneous across subjects since the group estimation is misleading (as previously discussed), and we want to observe the distribution of parameters across subjects.

[9] We will discuss these 12 subjects later in Section 6.

[10] $d \in \{0, \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}\}$, the optimal choice's chance of winning is $\frac{2}{3}$, $q_1 = 0.1$, $q_2 = 1 - q_1$, and for each d we have 100 simulated subjects.

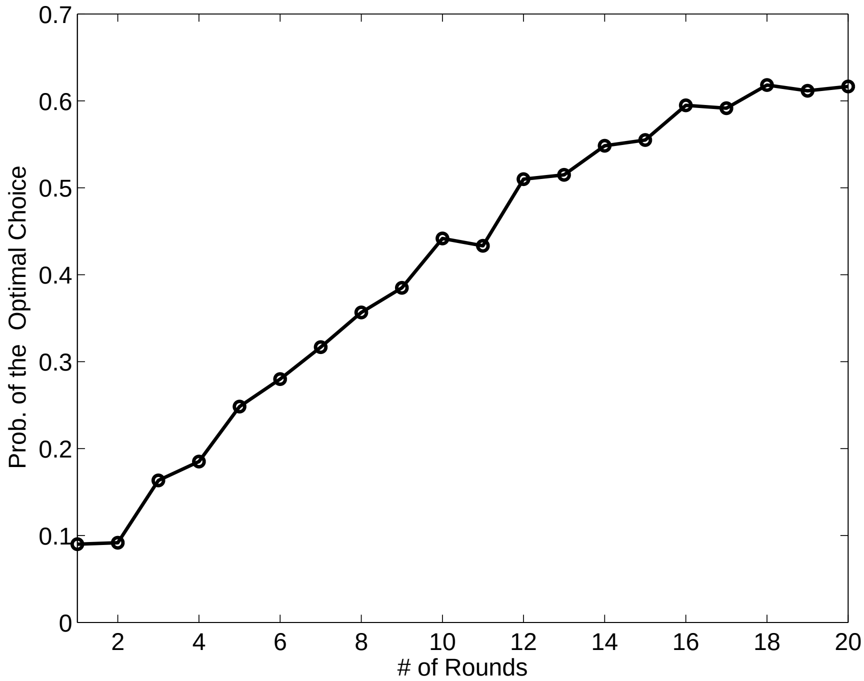

Figure 1: Average learning curve from a group of epiphany learners with different d.

Specifically, for a single subject, given a vector of parameters (in the EL model: (d, q_1, q_2, K); in the RL model: $(A^1(0), \phi, \lambda, K)$), both models provide us with a sequence of predicted probabilities of choosing 0, which is denoted by $\{\hat{p}_0(t)\}_{t=1}^{T}$, and we can thus define the Mean Square Deviation (MSD) by:

$$\sqrt{\frac{1}{T}\sum_{t=1}^{T}\{\hat{p}_0(t) - p_0(t)\}^2}$$

where $p_0(t)$ is subject's choice of probabilities on 0 in round t given in $s(t)$.

When estimating the parameters of both models, MSD will be the function we try to minimize, since a lower MSD indicates a better prediction. We searched over a range of parameter values[11] and performed the minimization procedure using the fmincon function in MATLAB, with many different starting points to avoid converging to a local minimum.

5.3.2. Model estimation results

For the EL model, the average of the estimated parameters are: $d = 0.41$, $q_1 = 0.098$, $q_2 = 0.95$, and $K = 0.52$. On the other hand, the RL model's estimation results are: $\lambda = 621.38$, $\phi = 0.28$, $A^1(0) = 438.28$, and $K = 1$. The median MSD from the EL model (0.0759) is significantly lower than the one from the RL model (0.1278) by the Wilcoxson signed-rank test (p-value < 0.001). In fact, only 14 out of 42 subjects were better fitted by the RL model. In order words, we identify 67% of our subjects as epiphany learners. This

[11] $\lambda \geq 0$, $0 \leq \phi, d, q_1, q_2 \leq 1$, $q_1 \leq q_2$, and $K \in \{0, 1, \ldots, 10\}$.

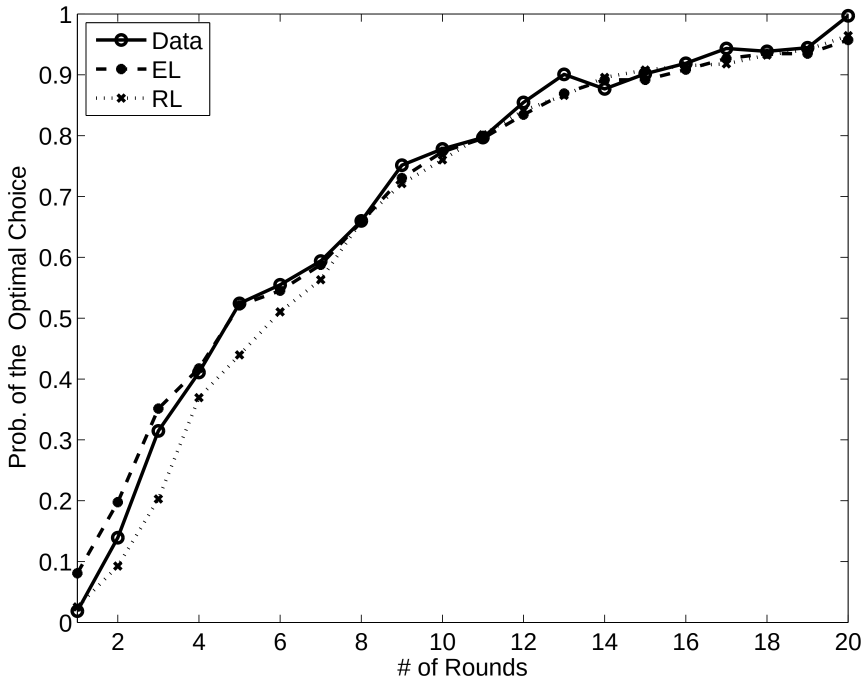

Figure 2: Data and models' predictions (Mean).

number is higher than the two previous studies (36% in Dufwenberg et al. (2010) and 35% in McKinney and Van Huyck), but if we include all 62 subjects into our analysis, then the number goes down to 45%. (see Table C.2 in the appendix for details). And compare to our previous study (Chen and Krajbich, 2017), where we found 38.9% of our subjects to be epiphany learners, this result seems to confirm that the EL proportion estimated from the commitment method was underestimated.

On the other hand, as shown in Figure 2, the average group choice behavior and predictions from both models are similar to each other, which again confirms that even if the individual learning behavior is better fitted by the EL model than the RL model, the average learning curve from the two models can be very similar. Therefore, we need to look at each individual subject's model fit results to understand why the EL model performs better than the RL model.

Figure 3 shows four representative examples of subjects who were significantly better fitted by the EL model. Note that the predictions from both models are similar in the first and last few rounds. This suggests that both models can capture the fact that in the first few rounds the decision-maker tends to not put any probability on 0, but after some cutoff point, she starts to put all probability on 0. However, in the middle rounds, the two models yield different predictions. The EL model is able to capture the sudden increase in the probability of choosing the optimal choice from 0 to 1, while the RL model needs more rounds to gradually catch up with that step. In other words, the RL model cannot capture the sudden burst in the learning curve, which we call "epiphany" in our EL model.

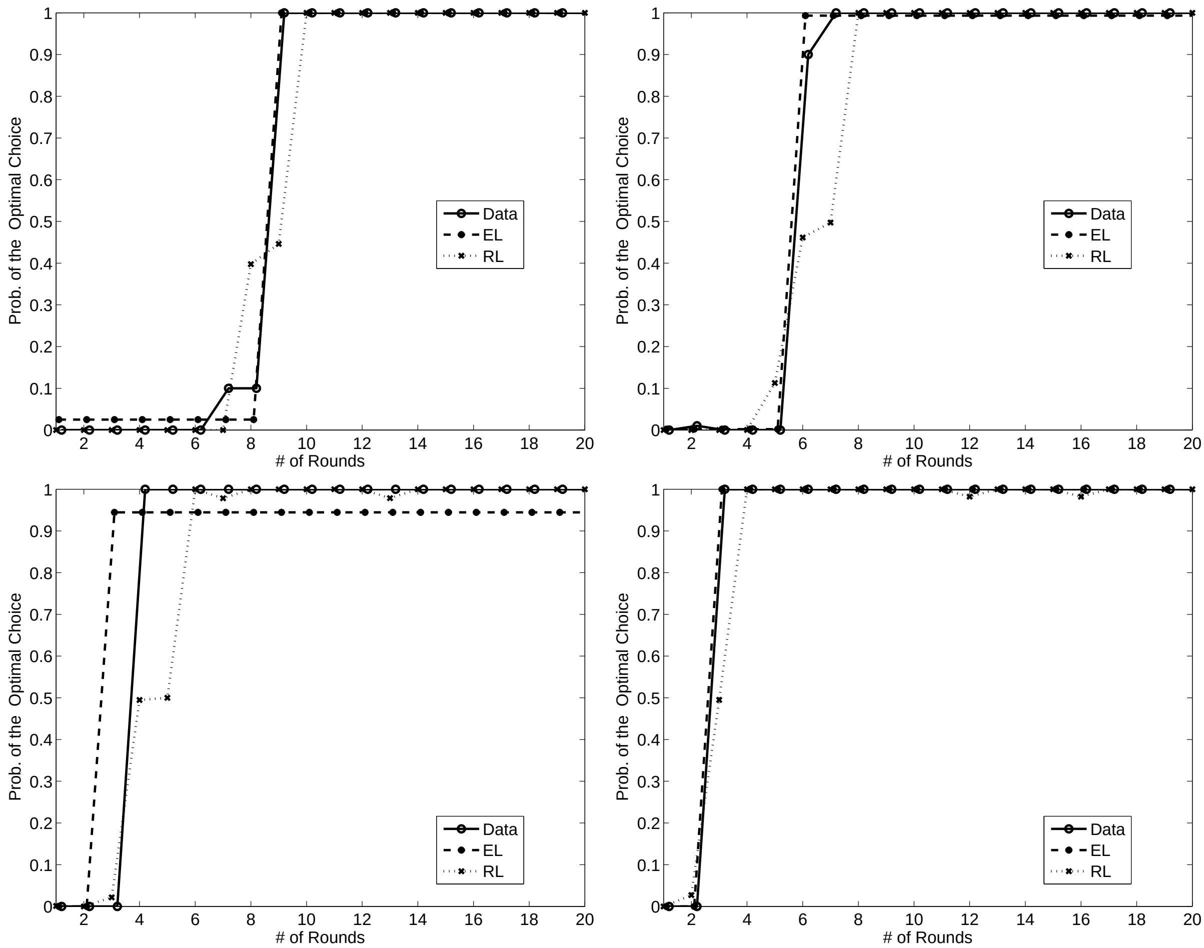

Figure 3: Data and models' predictions (Subject #4, 8, 14, 28, from left to bottom).

5.4. Distributions of the EL model estimation parameters

Now, let us look at the distributions of the four estimated parameters from the EL model. Figure 4 gives us the marginal distributions of all four parameters.

For d, we observe that 57.1% of our subjects have d greater than or equal to 1/3. Recall that a decision-maker with $d = 1/3$ would have an epiphany, if and only if the amount of evidence she receives from one side (positive or negative) exceeds the other side by three more occurrences. Therefore, for these subjects, the 2BC is relatively easy and thus they can figure it out just in a few rounds. However, for the rest of the subjects, the 2BC is not easy, and they need at least four rounds to realize that putting full weight on 0 is optimal.

For q_1, 69% of our subjects have q_1 less than 1/11. Because 1/11 is the probability density on any number (if the decision-maker uses a uniform random strategy), the result above suggests that most subjects believe 0 is worse than other numbers initially. This implies that they were far from figuring out the dominant strategy in round one.

For q_2, since we exclude those subjects who did not figure out the dominant strategy, its density is concentrated on 0.9 to 1 (as it should be).

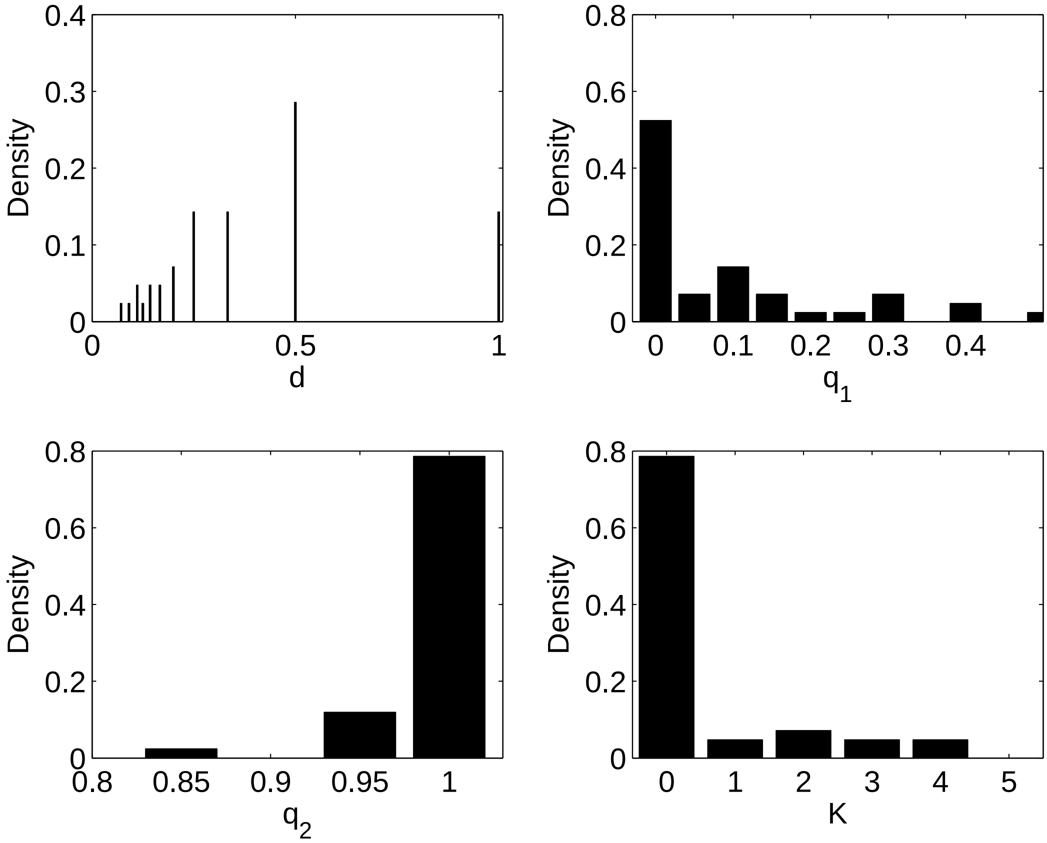

Figure 4: Distributions of the EL model parameters.

Lastly, for K, most subjects (79%) have an estimated K equal to 0. Recall that subjects treat $\{0, 1, 2, \ldots, K\}$ and $\{K+1, \ldots, 10\}$ as two choice partitions in the EL model. The first partition's success is viewed as evidence supporting 0, and vice versa. Thus, the distribution here suggests that most subjects only regard 0's successes as positive support for 0. They do not treat a success of one as positive evidence to support the idea that choosing 0 is the optimal strategy.

5.5. Robustness checks

We have shown that the EL model outperforms the RL model in fitting our experimental data. Here we present two alternative EL models for robustness test.

5.5.1. Random EL model

McKinney and Van Huyck (2013) use a structure change estimation method to estimate an epiphany point, and then test if the choice distribution before and after that point are significantly different. Although they do not propose a model of epiphany, their estimating method implies a random epiphany model.

Specifically, in this model, the epiphany moment follows a truncated Poisson distribution, with a mean of λ and a range of $[0, \bar{T}]$. The model predicts that subjects choose 0 with probability q_1 before the epiphany, and q_2 after the epiphany. Here λ, $\bar{T}$, q_1, and q_2 are parameters to be estimated[12]. Note that the epiphany moment is a random variable in this model. Thus we will take the expected value of the MSD as the MSD.

Following the same routine that we used to estimate the RL and EL model, we find that the average estimated parameters are $\lambda = 19.94$, $\bar{T} = 6.43$, $q_1 = 0.09$, and $q_2 = 0.92$, and the median MSD of this random epiphany model (0.094) is significantly higher than the median MSD from the EL model (0.0696) by the Wilcoxon signed-rank test, with p-value = 0.0012, but significantly lower than the median MSD from the RL model (0.1296), with p-value = 0.0016. Therefore, the random epiphany model performs worse than the EL model, but still outperforms the RL model. This result further confirms that a significant percentage of our subjects are epiphany learners.

5.5.2. Single-boundary EL model

Our EL model assumes that if enough negative evidence were aggregated, a decision-maker could have a negative epiphany. However, one can argue that an epiphany could only be positive because the optimal choice is unique. Here we estimate another EL model that assumes only the existence of the positive threshold, and test it with our data. Unlike our default EL model, here even if one's net evidence is lower than -1, she will keep accumulating evidence until a positive epiphany occurs. We found that the median MSD estimated from the single-boundary EL model (0.0759) is not significantly (Wilcoxon signed-rank test, p-value = 0.125) different from to the default EL model.

5.6. Predicting the EL

y = **EL**	**coef**	**std err**	**z**	p**-value**	**[0.025**	**0.975]**
pred_err	3.0047	1.128	2.663	0.008	0.793	5.217
dec_avg	-0.5688	0.639	-0.890	0.373	-1.821	0.683
dec_std	0.2908	0.503	0.578	0.563	-0.695	1.277
intercept	-0.4811	1.414	-0.340	0.734	-3.253	2.290

Table 1: Predicting if the subject is using EL. Logit regression results: **pred_err** is subject's decision (realized from the mixed strategy) minus the target number in the same round and **dec_avg / dec_std** is the expected value / standard deviation of her mixed strategy. Data are averaged over all rounds for a subject, excluding rounds that subjects put at least 950 black balls into 0, and only include the 42 subjects as in the model comparison exercise.

In the last part of the result, we sought to ask if we can predict whether a subject will learn the optimal strategy by EL or RL, using data from rounds before they have learned. We hypothesized that two objects can potentially helps: The first one is the choice prediction error (**pred_err**), which is defined as subject's decision (realized from the mixed strategy) minus the target number in the same round. A higher prediction error means a bigger mistake, which may trigger epiphany more easily. The second one is the mixed

[12] $0 < \lambda, \bar{T} < 20$ and $0 < q_1, q_2 < 1$.

strategy's property, including the expected value (**dec_avg**) and standard error (**dec_std**). For instance, a more disperse mixed strategy may suggest that subjects have no clue about what to choose at all and hence harder to get an epiphany. Therefore, we run the following logistic regression to confirm our hypothesis:

$$Pr(EL_i = 1 \mid X_i) = logit(intercept + \beta_1 * pred\_err_i + \beta_2 * dec\_avg_i + \beta_3 * dec\_std_i),$$

where EL_i is the dummy variable equals to 1 if subject i is using EL, and the 3 independent variables are data averaged over all rounds for the subject i, excluding rounds that subjects put at least 950 black balls into 0.

Table 1 summarizes the logistic regression result. We found that only the choice prediction error is positively (p-value = 0.008) predicting the EL, suggesting that a bigger mistake in early rounds will result in EL later on.

6. Discussions

6.1. Replacing the RL?

It is important to note that we are not proposing that people always use EL. There are many other situations where RL are certainly superior. However, our results demonstrate that in some settings (and for some people) EL explains behavior better than RL. In other words, EL is a good complement to the current learning literature. We believe with more data in the future, the EL model can be proven to be valuable in predicting behaviors in other games/decision-making problems as well.

6.2. Why not other reinforcement models?

We only compare the EL model to one kind of RL model, so why not other forms of the RL model? To answer this question, we have to consider the most distinct property of the EL model: the kinked learning curve with incubation.

As shown before, the learning curve of a EL looks like a step function, in which the kinked point happens at the epiphany moment, which is also called the S-shaped learning curve. Most RL models do not have this property, at least for the usual range of learning rate. More importantly, even if we push the learning rate in these models to some extreme value, and make these learning curves look like one from the EL model, the kinked point would often happens at a wrong time point. Thus, although we only take into consideration a specific type of RL model for the horse race here, what we really would argue is that the EL model can beat all other existing models with similar properties, unless we have a RL model in which the learning rate is not a constant, and depends on the timing for some reasons.

6.3. Twelve subjects who never learn the optimal strategy

Recall that we exclude 20 subjects in the data analysis since they are not the main focus of this paper. They either seem to know the optimal strategy in the beginning ($n = 8$), or did not figure out the optimal strategy even in the last round ($n = 12$). Here, we analyze the estimation results from the second group, but not the first group, since any model would fit the first group almost perfectly. These 12 subject's estimation results are in Table C.2, with shaded color, but no $*$ sign on the subject id. We found that the median MSD estimated from the EL model is 0.0732, which is not significantly (p-value $= 0.07$, Wilcoxon signed-rank test) higher than that from the RL model (0.0595). We suspect that these subjects did not have enough rounds to learn, and if they did, we would be able to conclude that they are reinforcement learners.

6.4. Generalized EL and the EL model for other games

Of course, it remains to be seen how well the EL will generalize to other settings. Although this is an empirical question that we cannot answer directly here, previous papers (Dufwenberg et al., 2010; McKinney and Van Huyck, 2013) have demonstrated epiphany-like learning in the game of 21 and Nim. Thus, it is reasonable to believe that EL will be applicable in other contexts. What remains to be tested is how we could know in advance which game is an EL game.

On the other hand, the EL model can be easily applied to any game with a unique optimal/dominant strategy, where the epiphany is clearly defined, with some adjustment to the model's assumption. For example, suppose the game we consider does not have a strategy space that can be aligned on a real number space like the 2BC game, then we need to restrict the EL model to have $K = 0$.

6.5. Why do we get a higher proportion of epiphany learners using the mixed strategy method?

As discussed in the introduction, by using the pure strategy method, the estimated EL might be RL and vice versa. In other words, the estimated proportion of epiphany learners could be under or overestimated. Here, we found that the proportion of epiphany learners is higher than what was found in previous literature, suggesting that more EL was previously misidentified as RL than the other way around. However, this conclusion must be taken very carefully, because previous studies' experimental design did not only differ from ours in one dimension. They either use different games (Dufwenberg et al., 2010; McKinney and Van Huyck, 2013) or add more features (like commitment, as in Chen and Krajbich, 2017) to the experiment. We need to test the EL estimation difference between the mixed and pure strategy method more directly to conclude.

Acknowledgments

We thank Shu-Yu Liu and Ally Wu for providing research assistance, the Pennsylvania State University for financial support, and Kalyan Chatterjee, James Tybout, Peter Newberry, Richard Carlson, Ian Krajbich, Katie Coffman, Paul Healy, Joseph Tao-yi Wang, Stephanie Smith, Rachael Gwinn and Arkady Konovalov for helpful comments.

References

Arifovic, J., McKelvey, R. D., Pevnitskaya, S., 2006. An initial implementation of the turing tournament to learning in repeated two-person games. Games and Economic Behavior 57 (1), 93–122.

Arthur, W. B., 1991. Designing economic agents that act like human agents: A behavioral approach to bounded rationality. American Economic Review 81 (2), 353–359.

Biele, G., Erev, I., Ert, E., 2009. Learning, risk attitude and hot stoves in restless bandit problems. Journal of Mathematical Psychology 53 (3), 155–167.

Börgers, T., Sarin, R., 2000. Naive reinforcement learning with endogenous aspirations. International Economic Review 41 (4), 921–950.

Bower, G. H., 1961. Application of a model to paired-associate learning. Psychometrika 26 (3), 255–280.

Camerer, C., Ho, T.-H., 1999. Experience-weighted attraction learning in normal form games. Econometrica 67 (4), 827–874.

Camerer, C. F., Ho, T.-H., Chong, J.-K., 2004. A cognitive hierarchy model of games. The Quarterly Journal of Economics 119 (3), 861–898.

Chen, W., Liu, S.-Y., Chen, C.-H., Lee, Y.-S., 2011. Bounded memory, inertia, sampling and weighting model for market entry games. Games 2 (1), 187–199.

Chen, W. J., Krajbich, I., 2017. Computational modeling of epiphany learning. Proceedings of the National Academy of Sciences 114 (18), 4637–4642.

Chou, E., McConnell, M., Nagel, R., Plott, C. R., 2009. The control of game form recognition in experiments: understanding dominant strategy failures in a simple two person guessing game. Experimental Economics 12 (2), 159–179.

Cross, J. G., 1973. A stochastic learning model of economic behavior. The Quarterly Journal of Economics 87 (2), 239–266.

Dodds, R., Ward, T., Smith, S., 2003. A review of the experimental literature on incubation in problem solving and creativity. Vol. 3. Hampton Press Cresskill, NJ.

Dufwenberg, Martin, Sundaram, Ramya, Butler, D. J., August 2010. Epiphany in the game of 21. Journal of Economic Behavior & Organization 75 (2), 132–143.

Estes, W. K., 1956. The problem of inference from curves based on group data. Psychological Bulletin 53 (2), 134–140.

Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10 (2), 171–178.

Gill, D., Prowse, V., 2016. Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis. Journal of Political Economy 124 (6), 1619–1676.

Grosskopf, B., Nagel, R., January 2008. The two-person beauty contest. Games and Economic Behavior 62 (1), 93–99.

Grosskopf, B., Nagel, R., 2009. Unraveling in two-person guessing games, working Paper.

Hanaki, N., Kirman, A., Pezanis-Christou, P., 2018. Observational and reinforcement pattern-learning: An exploratory study. European Economic Review 104, 1–21.

Hélie, S., Sun, R., 2010. Incubation, insight, and creative problem solving: a unified theory and a connectionist model. Psychological Review 117 (3), 994–1024.

Ho, T.-H., Camerer, C., Weigelt, K., 1998. Iterated dominance and iterated best response in experimental "p-beauty contests". American Economic Review 88 (4), 947–969.

Jausovec, N., Bakracevic, K., 1995. What can heart rate tell us about the creative process? Creativity Research Journal 8 (1), 11–24.

Krajbich, I., Oud, B., Fehr, E., 2014. Benefits of neuroeconomic modeling: New policy interventions and predictors of preference. American Economic Review 104 (5), 501–506.

McKinney, C. N., Van Huyck, J. B., 2013. Eureka learning: Heuristics and response time in perfect information games. Games and Economic Behavior 79, 223–232.

Metcalfe, J., 1986. Premonitions of insight predict impending error. Journal of Experimental Psychology: Learning, Memory, and Cognition 12 (4), 623.

Metcalfe, J., Wiebe, D., 1987. Intuition in insight and noninsight problem solving. Memory & Cognition 15 (3), 238–246.

Mookherjee, D., Sopher, B., 1994. Learning behavior in an experimental matching pennies game. Games and Economic Behavior 7 (1), 62–91.

Murre, J. M., 2014. S-shaped learning curves. Psychonomic Bulletin & Review 21 (2), 344–356.

Nagel, R., 1995. Unraveling in guessing games: An experimental study. American Economic Review 85(5), 1313–1326.

Nagel, R., Bühren, C., Frank, B., 2016. Inspired and inspiring: Hervé moulin and the discovery of the beauty contest game. Mathematical Social Sciences.

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., Gold, J. I., 2012. Rational regulation of learning dynamics by pupil-linked arousal systems. Nature Neuroscience 15 (7), 1040–1046.

Pearce, J. M., 1987. A model for stimulus generalization in pavlovian conditioning. Psychological review 94 (1), 61.

Pols, A., 2002. Insight problem solving, unpublished doctoral dissertation. University of Utrecht, Utrecht, the Netherlands.

Ratcliff, R., 1978. A theory of memory retrieval. Psychological Review 85 (2), 59.

Roth, A. E., Erev, I., 1995. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. Games and Economic Behavior 8 (1), 164–212.

Salmon, T. C., 2001. An evaluation of econometric models of adaptive learning. Econometrica 69 (6), 1597–1628.

Wallas, G., 1926. The art of thought. London.

Appendix A. Models

Appendix A.1. EL model

The decision-maker plays the 2BC for T rounds, and receives feedback about the target number and whether she has won the current round at the end of each round. She has some initial belief about which strategy -possibly a mixed one- is optimal. She will only change her strategy if she accumulates enough evidence against her initial belief. We will call this moment that she changes her strategy the epiphany moment.

To define the evidence in favor of the belief that playing 0 is optimal (hereafter evidence), we first define $x(t)$ as the decision-maker's choice of integer at round t, and two sets of numbers representing "small numbers" and "large numbers": $X_0 \equiv \{0, 1, 2, \ldots, K\}$ and $X_1 \equiv \{K+1, \ldots, 10\}$, where K is a parameter in the model. Here numbers in X_0 are those that support the belief that playing 0 is optimal when these numbers lead to a win, and X_1 supports the belief that another number in X_1 is optimal when these numbers lead to a win.

The evidence ($ev(t)$) accumulated up to round t is defined as:

$$ev(t) = ev(t-1) + d \times (-1)^{I_t},$$

where $I_t = 1$ if $x(t) \notin X_0$ and the result is a win (or tie), or that $x(t) \in X_0$ and the result is a loss. $I_t = 0$ otherwise, and the initial evidence is $ev(0) = 0$. Here d is a parameter in the model.

Furthermore, to specify the model's prediction, we say that a positive (negative) epiphany occurs at round t^*, if and only if $ev(t^*) \geq 1$ (≤ -1). A decision-maker chooses 0 with probability q_1 initially[13], and after each positive epiphany, she starts to choose 0 with probability q_2, after each negative epiphany, she starts to choose another number $x' \neq 0$ with probability q_2, and thus choose all other numbers including 0 with probability $(1-q_2)/10$. Here, q_1 and q_2 are parameters in the model, and the threshold 1 is set arbitrarily due to the arbitrary scale of $ev(t)$.

To sum up, the evidence increase by d units if small numbers are guessed and the result is a win (or tie), or if large numbers are guessed and the result is a loss, and decrease by d units otherwise. These two forces compete against each other, and if one is much stronger ($1/d$ more) than the other, a corresponding epiphany will occur, and the decision-maker's belief/behavior will change accordingly.

There are two key assumptions in the EL model. The first one is the existence of both positive and negative epiphany threshold. We borrow this assumption from the standard DDM (Ratcliff, 1978; Krajbich et al., 2014), however, as shown in the section 5.5.2, relaxing this assumption does not change our results significantly.

The second assumption is that the success of numbers near 0 (from 0 to K) would help the epiphany. This comes from the generalization in conditioning literature in psychology (Pearce, 1987). As we will show later, most subjects are estimated to have $K = 0$, suggesting that this assumption is not crucial either. We still use this assumption to make the EL model as general as possible.

[13] We only make a prediction about the probability that a decision-maker chooses to allocate on 0, because 0 is the weakly dominant strategy here.

Appendix A.2. RL model

Here we employ the RL model in Salmon (2001). In order to keep the two models in our paper compatible, we also group choices into X_0 and X_1 separately for RL, the same as in the EL model.

In the beginning ($t = 0$), the agent has some prior attractions $A^j(0)$ for X_j, $j \in \{0, 1\}$, where $A^0(0)$ is normalized to be 0, and $A^j(t)$ is updated by the following equation:

$$A^j(t) = \phi A^j(t-1) + \mathbb{1}(x(t) \in X_j)\pi(t)$$

where $\phi \in (0, 1]$ is the discount factor for the attraction, $\pi(t)$ is the payoff received in round t, and $\mathbb{1}(\cdot)$ is the indicator function. In other words, the agent starts with the initial attraction, and in the end of each round, these attractions are discounted, and if a choice in X_j is realized, the associated payoff would also be added to the attraction for X_j. Finally, the predicted choice probability on choosing 0 at round t is given by:

$$\hat{p}_0(t+1) = \frac{e^{\lambda A^0(t)}}{e^{\lambda A^0(t)} + e^{\lambda A^1(t)}}.$$

Note that in this model, three more parameters ($A^1(0)$, ϕ, λ) need to be estimated in addition to K.

Appendix B. Experimental design for the database experiment

In its traditional form, the 2BC is a two-person strategic game, not a decision problem. Therefore, it is possible for a subject to learn that playing 0 is optimal because she observes that her opponent's choice sequence is converging to 0 (Grosskopf and Nagel, 2009). To avoid this type of learning channel, we transform the 2BC into a decision problem so that our subjects are not able to learn the optimal strategy from one another. Specifically, we let our subjects play against a fixed database. An initial group of subjects repeatedly played the 2BC game ($p = 0.9$) against each other for 20 rounds, with random rematching and feedback that included the target number and game result at the end of each round. Their choice data became the database. A separate group of subjects in the formal experiment played against this database[14].

Note that both groups of subjects were asked to provide a mixed strategy in each round, using the graphic interface displayed in Figure B.5. Subjects were asked to distribute 1000 black balls into 11 positions corresponding to the 11 possible choices (0 to 10). They were only allowed to hit the confirm key if the total number of black balls in all 11 positions summed to exactly 1000, and they had a calculator available in the bottom-right corner in case they needed some help with the algebra. They were told in the written instructions that one of these black balls would be picked, and their choice would depend on that black ball's position. For example, in the Figure B.5, all 1000 black balls were evenly distributed between 0 and 1. This is effectively a mixed strategy that with 50% chance of choosing 0 and 50% chance of choosing 1.

[14] We are fully aware of the fact that this setting changes the game, but we believe that allowing subjects to learn from their opponent's choice sequence would more severely contaminate our data.

Figure B.5: Screenshot from the choice screen. Traditional Chinese was used in the experiment. Here the English translation is provided next to the original text.

Appendix C. Individual Estimation Results

	EL Model							RL Model					Random Epiphany Model				
Subject	Parameters				EM	EM		Parameters					Parameters				
id	d	q_1	q_2	K	(+)	(−)	MSD	λ	ϕ	$A1_0$	K	MSD	λ	q_1	q_2	$\bar{T}$	MSD
1*	0.13	1.00	1.00	0	8	–	0.00	11.00	0.55	-30.00	0	0.00	19.79	1.00	1.00	10	0.00
2	0.20	0.01	0.94	0	7	–	0.10	999.99	0.10	999.98	0	0.08	20.00	0.06	0.95	8	0.15
3*	0.00	1.00	1.00	0	–	–	0.00	11.00	0.55	-30.00	0	0.00	3.02	0.50	1.00	0	0.00
4	0.25	0.03	1.00	1	8	–	0.03	92.23	0.52	999.44	2	0.14	20.00	0.02	0.99	8	0.11
5	0.33	0.00	1.00	0	3	–	0.00	998.34	0.00	998.34	0	0.16	19.98	0.00	1.00	3	0.03
6	0.00	0.00	1.00	0	–	–	0.00	468.44	0.46	22.25	0	0.00	3.12	0.52	0.00	0	0.00
7	0.50	0.20	1.00	0	2	–	0.06	980.32	0.00	0.59	0	0.11	20.00	0.20	1.00	2	0.07
8	0.33	0.00	0.99	0	5	–	0.02	999.99	0.04	999.99	0	0.15	20.00	0.00	0.99	5	0.07
9	0.50	0.15	1.00	0	6	–	0.08	125.85	0.31	999.86	2	0.14	20.00	0.15	0.99	6	0.11
10	0.14	0.02	0.98	2	7	–	0.02	0.46	0.58	761.89	2	0.02	20.00	0.02	0.97	7	0.10
11	0.00	0.00	1.00	0	–	–	0.00	314.92	0.57	7.52	0	0.00	3.12	0.52	0.00	0	0.00
12	1.00	0.00	0.95	3	1	–	0.22	447.01	0.10	99.10	3	0.13	19.96	0.00	1.00	2	0.02
13	0.11	0.32	1.00	0	11	–	0.19	0.06	1.00	32.53	3	0.18	20.00	0.30	0.96	11	0.20
14	0.50	0.00	0.94	3	2	–	0.22	999.80	0.00	19.80	0	0.16	19.98	0.00	1.00	3	0.03
15	1.00	0.09	0.99	0	1	–	0.00	998.45	0.02	0.00	0	0.11	18.31	0.09	0.99	1	0.01
16	1.00	0.00	1.00	4	3	1	0.00	998.53	0.00	998.53	0	0.11	19.98	0.00	1.00	3	0.03
17	1.00	0.00	1.00	0	1	–	0.00	998.23	0.00	0.00	0	0.11	19.92	0.00	1.00	1	0.01
18	0.17	0.00	0.98	0	8	–	0.04	999.96	0.06	970.50	0	0.17	20.00	0.00	0.97	8	0.12
19	0.50	0.00	0.16	0	6	2	0.23	0.00	0.98	996.24	5	0.24	19.99	0.02	0.16	6	0.23
20	0.13	0.00	0.87	0	12	–	0.21	0.14	0.83	164.77	2	0.25	20.00	0.05	0.88	14	0.27
21	0.50	0.00	1.00	0	2	–	0.00	999.86	0.00	9.26	0	0.11	19.96	0.00	1.00	2	0.02
22	0.17	0.07	0.62	0	12	–	0.14	0.05	0.67	999.99	0	0.08	20.00	0.03	0.57	11	0.15
23	0.33	0.00	0.12	0	7	–	0.11	0.08	0.85	37.73	0	0.11	20.00	0.00	0.12	7	0.11
24*	0.00	1.00	1.00	0	–	–	0.00	11.00	0.55	-30.00	0	0.00	3.02	0.50	1.00	0	0.00
25	0.50	0.00	1.00	0	8	2	0.04	999.97	0.02	0.59	0	0.16	20.00	0.05	0.99	8	0.11
26	1.00	0.00	1.00	0	1	–	0.00	999.99	0.02	0.00	0	0.16	19.92	0.00	1.00	1	0.01
27	0.50	0.10	1.00	0	6	–	0.12	846.36	0.23	999.97	5	0.13	20.00	0.00	0.97	5	0.13
28	0.50	0.00	1.00	0	2	–	0.00	999.86	0.00	9.26	0	0.11	19.96	0.00	1.00	2	0.02
29	0.00	0.00	1.00	0	–	–	0.00	2.99	1.00	562.50	4	0.00	3.12	0.52	0.00	0	0.00
30	0.20	0.08	1.00	0	19	–	0.22	499.64	0.85	0.01	1	0.17	20.00	0.03	0.51	18	0.23
31	0.33	0.08	1.00	0	13	–	0.21	999.99	0.02	822.79	0	0.12	20.00	0.01	0.99	12	0.17
32	0.14	0.14	0.50	0	19	–	0.08	0.04	0.82	182.87	0	0.06	20.00	0.11	0.29	17	0.09

33	0.33	0.30	1.00	0	3	–	0.08	516.07	0.05	0.02	0	0.11	20.00	0.29	1.00	3	0.09
34	0.50	0.00	1.00	0	2	–	0.00	999.86	0.00	9.25	0	0.11	19.96	0.00	1.00	2	0.02
35*	0.00	1.00	1.00	0	–	–	0.00	11.00	0.55	-30.00	0	0.00	3.02	0.50	1.00	0	0.00
36	0.10	0.22	0.42	1	14	–	0.08	0.00	0.90	763.31	0	0.06	20.00	0.16	0.35	8	0.08
37*	0.00	1.00	1.00	0	–	–	0.00	11.00	0.55	-30.00	0	0.00	3.02	0.50	1.00	0	0.00
38	0.50	0.00	0.47	2	4	–	0.09	0.02	0.52	1000.00	0	0.06	20.00	0.02	0.50	5	0.04
39	0.25	0.06	0.96	0	10	–	0.09	165.26	0.57	999.97	1	0.08	20.00	0.08	0.96	11	0.15
40	0.07	0.86	1.00	0	14	–	0.24	1.00	0.55	-30.00	0	0.26	19.99	1.00	0.85	7	0.24
41	0.33	0.00	1.00	0	7	–	0.00	999.99	0.02	0.91	0	0.16	19.96	0.00	1.00	7	0.08
42*	0.00	1.00	1.00	0	–	–	0.00	11.00	0.55	-30.00	0	0.00	3.02	0.50	1.00	0	0.00
43	0.14	0.15	0.99	0	9	–	0.11	0.37	1.00	8.16	1	0.18	20.00	0.13	0.97	9	0.14
44	0.50	0.39	1.00	0	2	–	0.06	997.80	0.00	0.00	0	0.11	20.00	0.39	1.00	2	0.07
45	0.25	0.13	1.00	0	4	–	0.02	999.98	0.02	171.11	0	0.12	20.00	0.12	1.00	4	0.05
46	0.20	0.08	1.00	0	9	–	0.10	0.44	1.00	27.53	3	0.08	20.00	0.06	0.98	9	0.13
47	0.50	0.05	1.00	0	2	–	0.01	999.99	0.07	0.03	0	0.12	20.00	0.05	1.00	2	0.03
48	0.50	0.25	1.00	0	4	–	0.19	164.44	0.10	936.20	4	0.13	19.98	0.00	1.00	3	0.03
49	0.17	0.00	0.25	0	18	–	0.08	2.05	0.97	56.80	2	0.00	20.00	0.00	0.14	19	0.10
50	1.00	0.00	0.95	2	1	–	0.22	266.89	0.01	1000.00	2	0.11	19.96	0.00	1.00	2	0.02
51	0.25	0.03	0.08	2	14	–	0.07	6.35	0.70	218.66	2	0.07	19.97	0.15	0.04	1	0.07
52	0.33	0.31	1.00	0	15	–	0.14	0.03	1.00	72.49	4	0.13	20.00	0.29	0.89	15	0.18
53	0.17	0.00	0.75	1	16	–	0.19	0.06	0.78	996.92	2	0.27	20.00	0.00	0.54	16	0.27
54*	0.00	1.00	1.00	0	–	–	0.00	11.00	0.55	-30.00	0	0.00	3.02	0.50	1.00	0	0.00
55	0.25	0.00	1.00	0	4	–	0.00	999.59	0.01	999.59	0	0.11	19.91	0.00	1.00	4	0.05
56	0.09	0.00	0.33	2	17	–	0.18	998.69	0.02	911.67	0	0.16	20.00	0.00	0.29	19	0.19
57	0.33	0.05	0.15	0	7	–	0.06	0.01	0.95	298.73	0	0.06	20.00	0.04	0.15	7	0.06
58	0.25	0.00	1.00	0	4	–	0.00	999.44	0.01	999.44	0	0.16	19.91	0.00	1.00	4	0.05
59*	0.00	1.00	1.00	0	–	–	0.00	11.00	0.55	-30.00	0	0.00	3.02	0.50	1.00	0	0.00
60	0.11	0.08	1.00	4	11	–	0.07	3.07	0.92	1.26	0	0.06	20.00	0.07	0.97	11	0.16
61	0.33	0.01	1.00	5	–	11	0.02	21.82	1.00	10.22	3	0.02	19.97	0.03	0.00	3	0.02
62	0.25	0.38	1.00	0	4	–	0.19	999.82	0.02	0.00	0	0.22	20.00	0.38	0.99	4	0.20

Table C.2: Individual estimation results: Epiphany Moment (EM) indicate the estimated timing of positive/negative (+/−) epiphany. For all subjects, each type of epiphany happened at most once. Shaded rows represent subjects who were not included in the data analysis, and ∗ indicates that the subject always puts at least 95% of the probability on 0.

Appendix D. Experimental Instruction

This experiment has 20 rounds. Every round asks you to guess a number between 0 and 10. The computer will calculate the average number out of the guesses in the same group and multiply it by 0.9 to get the target number. The person whose guess is the closest to the target number will win the round and $10 NTD. If there are two winners, the computer will choose a winner randomly. For example, if everyone in the group chooses 10 as their number, then the target number will be 9. Therefore, everyone's number is as close as everyone else's. At this moment, the computer will randomly pick a winner.

In this experiment, the way you guess a number is by distributing 1000 black balls. After you finished your distribution, the computer will randomly pick a black ball and the number you assign to it will be your guess. For example, if you put 600 black balls on number 5, then other 400 black balls will be evenly distributed on other 10 numbers. If the computer picks the ball out of the 600 black balls, then your guess will be 5. Please be sure that you need to use up every black ball and each of them cannot be divided (e.g. you cannot put 0.5 ball on a number.) Every number will have 0 to 1000 black balls. If you need to calculate how many balls you have distributed, there will be a calculator icon at the bottom right-hand corner. After you click on it, there will be a calculator for you to use.

Before the experiment starts, you'll have a practice round. In this round, every participant will be in the same group. This is only for you to be familiar with how the experiment is going to proceed and the interface. This practice round has nothing to do with your cash reward. After the practice round, the host will announce that "the experiment officially starts" Then we'll enter the official experiment.

In the official experiment, each group has two people. However, the other member in your group is not anyone in the room; he or she is randomly picked from NTU students who participated in this experiment in the past (the computer will pick a new member every round.) This group of NTU student took exactly the same experiment, but when they entered the experiment, their partner was randomly picked from other participants in the room at the time. Every round, the computer will randomly pick the number from the decisions these students made in their first 15 rounds. Your reward will be calculated accordingly. Those NTU students will not receive any extra reward or be charged an extra fee because of your choice.