國立臺灣大學社會科學院經濟學系
碩士論文
Department of Economics
College of Social Sciences
National Taiwan University
Master Thesis

用經濟學實驗研究貝氏修正偏誤差異與同溫層效果
Asymmetric Failure of Bayesian Updating and the Echo Chamber Effect: An Experimental Study

鄭淳厚
Chun-Hou Cheng

指導教授：
王道一 博士
狄萊 博士

Advisor:
Joseph Tao-yi Wang, Ph.D
Patrick DeJarnette, Ph.D

中華民國 109 年 8 月
August, 2020

doi:10.6342/NTU202002331

國立臺灣大學碩士學位論文
口試委員會審定書

Asymmetric Failure of Bayesian Updating and the Echo Chamber Effect: An Experimental Study

本論文係鄭淳厚君（學號 R05323035）在國立臺灣大學經濟學系完成之碩士學位論文，於民國 109 年 3 月 23 日承下列考試委員審查通過及口試及格，特此證明

口試委員：

（指導教授）

doi:10.6342/NTU202002331

誌謝

就讀經研所的這段時間，很感謝 Joseph 和 Patrick 的指導，兩位都大幅精練我思考的能力，並且熱心地提供我生涯規劃的建議。從剛入學擔任研究助理，到最後修改論文的過程，Joseph 一直很有耐心的和我討論，我永遠不會忘記初三那天一早竟然還願意花時間和我討論。Patrick 在本鄉三丁目鼓勵我繼續從事學術研究的對談也令我畢生難忘。

入學一個月左右時進入 TASSEL 實驗室，這是在我研究所中做過最正確的決定之一，在實驗室認識了許多優秀的學長姊——帶我熟悉實驗流程的孟謙，一起到國外參加研討會的孟章和建勳，還有回國擔任教職的 Josie、依依學姊、陳暐學長在每次報告及討論時給我非常有幫助的建議，協助實驗及提供我求職方向的曉芳姊。實驗室中也有一起努力的朋友——在做研究的路上總是有共鳴的呂越和昱翔，給我研究和修課建議的伯軒，當然還有陸陸續續一起在實驗室合作的眾多夥伴。

另外，我想特別感謝對我影響很深的幾位老師。大學時因為修了古慧雯老師的個體經濟學和蔡崇聖老師的賽局與訊息，開啟我經濟學的興趣，並決定報考經研所。研究所時修了黃貞穎老師的神經經濟學，讓我對做研究有更深入的理解。還有願意給我機會擔任統計與計量助教的李宗穎老師，讓我瞭解如何把自己的知識清楚的傳達給他人。

最後，謝謝我的父母和家人願意讓我嘗試學術這條路，並且給我不同面向的支持。由衷感謝在完成這個學位的路上曾經對我有任何或大或小幫助的人。

doi:10.6342/NTU202002331

用經濟學實驗研究貝氏修正偏誤差異與同溫層效果

中文摘要

人們如何處理與原先信念衝突但可能有幫助的資訊？我們用經濟學實驗來研究人在面對這類資訊時如何進行貝氏修正。實驗中電腦將每個人獨立地隨機指定到兩個盒子之一、從中抽出一顆球，然後獲得另一個人所申報的新資訊，但他只有一半的機會跟自己是同一個盒子。我們發現當新資訊與自己原先的資訊相衝突時，人們更新信念的結果會更偏離貝氏更新的預測，表示他們將新資訊過度地歸因於與自己不同的盒子，而且即使考慮人們認定他人有可能胡亂申報，這個效果仍然存在。因此，當新資訊與原先信念衝突時，人們會有將之歸因於不相干的傾向，進而忽略新資訊、不願更新信念，形成穩定的同溫層均衡。

JEL 分類碼：C44, D91, C91

關鍵字：貝氏定理，貝氏修正偏誤，不對稱更新，同溫層效應，兩極化效應

doi:10.6342/NTU202002331

Abstract

How do individuals update beliefs from contradicting information that could be potentially irrelevant? We conducted a laboratory experiment to investigate the ability to process such information, in which each subject independently draws a ball from one of two digital urns and receives information reported by another subject who may or may not have drawn from the same urn. We document evidence that subjects who receive new conflicting information deviate more from Bayesian updating, indicating that subjects overly believe the new information is irrelevant. This effect is robust even allowing subjects to perceive others reporting randomly. This pattern of attributing conflicting information as irrelevant may form the foundation of stable echo chambers or equilibria where additional information has no effect on beliefs.

JEL codes: C44, D91, C91

Keywords: Bayes' rule; polarization; belief-updating; asymmetric processing; biased interpretation

doi:10.6342/NTU202002331

目　錄

口試委員審定書…………………………………………………………… i

誌謝………………………………………………………………………… ii

中文摘要………………………………………………………………… i ii

英文摘要………………………………………………………………… iv

1. Introduction………………………………………………………… 1

2. Experimental Design……………………………………………... 5

2.1 Design Details……………………………………………... 5

2.2 Belief Eliciation…………………………………………… 6

2.3 Experimental Procedures…………………………………….. 7

2.4 Bayesian Probability Predictions…………………………….. 9

3. Results……………………………………………………………..16

3.1 Adherence to Bayesian Updating…………………………….. .16

3.2 The Echo Chamber……………………………………………..19

3.3 Inferred Prior Beliefs of Other's Information………………… 25

4. Conclusion………………………………………………………… 33

References…………………………………………………………… 34

Appendix: First Phase Belief……………………………………………36

Appendix: Second Phase Raw Data…………………………………… .36

Appendix: Alternative Experimental Designs…………………………...37

Appendix: Experimental Instructions……………………………………41

doi:10.6342/NTU202002331

圖 目 錄

Figure 1: Two-stage Menu of Lottery Choices……………………………8

Figure 2: Elicited Beliefs in the First Phase…………………………… 17

Figure 3: Elicited Beliefs Distribution in the Second Phase……………19

Figure 4: Type of Behavior (Irrelevant Urn) …………………………...21

Figure 5: Elicited Beliefs of the Irrelevant Urn…………………………23

Figure 6: Fully Dissociate …………………………………………………24

Figure 7: Model of the Information Sources……………………………27

Figure 8: Information Sources Distributions……………………………29

Figure 9: Beliefs Aligned with 45 Degree Line in the Irrelevant Urn…..36

Figure 10: Elicited Beliefs in the Second Phase………………………...37

Figure 11: Screen Shot of the First Version Experiment………………..38

Figure 12: Screen Shot of the Third Version Experiment………………40

doi:10.6342/NTU202002331

表　目　錄

Table 1: Percentage of Theory-consistent Choices Under Different Error Margins…………………………………………………………………… 17

Table 2: Type of Behavior (Irrelevant Urn) …………………………… 21

Table 3: Attribution of the Information…………………………………22

Table 4: Individual Type Transition…………………………………….30

Table 5: Independent Source Beliefs……………………………………31

Table 6: Individual Source Beliefs………………………………………32

doi:10.6342/NTU202002331

1 Introduction

Information processing plays an important role in many life decisions, but the same information may not always be interpreted the same way. For example, despite scientific consensus on climate change, individuals may have different interpretations of the likely impact that persist for long periods of time (Kahan et al., 2011, 2012; Fryer Jr et al., 2019). Receiving a bad grade in math may lead some to pursue non-STEM degrees, while others may see it as a challenge to persist. Some may even believe the earth is flat despite apparent evidence. Indeed, personal experience and motivation may play a role in individual interpretation.

When information could be potentially irrelevant, individuals may be overly inclined to discount it entirely when it contradicts their prior beliefs. For example, if one believed there was a chance new research was driven by political or commercial interests, it may be overly easy to discount this as untrustworthy or irrelevant. In contrast, when this information conforms to our pre-existing beliefs, it may be much harder to account for this possibility. Since Bayes' Rule implies we should still update our posteriors according to the new information in both cases, ignoring such information because it may be irrelevant results in incorrect beliefs persisting for much longer.

In this paper, we examine a laboratory experiment investigating people's ability to process new information, and study the difference in belief updating when information aligns with or is against the prior belief. Specifically, we consider a two step procedure, in which subjects first independently draw information from one of two urns, and use this information to update beliefs about the state of that urn. Then,

doi:10.6342/NTU202002331

each subject learns the stated belief of another randomly chosen subject for the urn that he or she drew from. However, the second subject's urn may or may not be the same as the first subject.

In the face of conflicting information, a subject should correctly infer that the other subject is more likely to have drawn from a different urn. However, this does not mean that the other subject could not have drawn from the same urn, just that it is less likely. However, even in a neutral context, we document that subjects appear to overestimate the probability that conflicting information comes from a different (and hence, irrelevant) urn. Thus, in our neutral setting subjects underestimate the chance that conflicting information implies that their own initial information was misleading. Conversely, in the face of confirming information, subjects overestimate the probability that this information comes from the same urn, and underestimate the chance that seemingly confirming information may be irrelevant (coming from a different urn).

Fryer Jr et al. (2019) introduce a model to depict why polarization in people's beliefs would occur in many settings where information is open to interpretation. An important theoretical prediction from this paper is that polarization increases when people interpret an ambiguous signal as a certain signal for a particular state based on their current beliefs. Their online Amazon Mechanical Turk experiments show that when subjects observe a sequence of information, they indeed form biased interpretation of evidence in the face of ambiguous ones and results in polarization in issues like climate change and death penalty.

In contrast, we provide three main contributions. First, the polarizing beliefs in

doi:10.6342/NTU202002331

Fryer Jr et al. (2019) stem from ambiguous information that is incorrectly inferred to be informative. In our paper, we explore how individuals incorporate information that is contradictory to their current beliefs, rather than how they misinterpret non-informative signals. Thus, even in purely informative spaces, we show improper Bayesian updating. Secondly, in our experiment, we explore a politically neutral context with objective outcomes, as opposed to the politically charged context with subjective outcomes (as the scale of interpretation in Fryer Jr et al. (2019) may be itself differently interpreted based on prior beliefs). Lastly, we provide evidence that contradictory information is not misinterpreted as consistent with beliefs, but is instead tends to be misattributed as irrelevant. Collectively, these results can potentially explain why, despite the general scientific consensus on climate change, individuals may form beliefs that cause them to ignore this information. In other words, given the relative paucity of ambiguous information in climate change, it may be that individuals instead infer that unambiguous contradictory information is instead from an untrustworthy or irrelevant source.

Failure of Bayesian updating is documented in several papers, including Tversky and Kahneman (1973) and Grether (1980). They find that subjects ignore base-rate information contrary to Bayes rule, resulting in representativeness heuristic. Holt and Smith (2009) show that subjects tend to over/underweight low/high prior probabilities. Compare to cognitive incompetence to perform Bayes rule, recent studies focus on asymmetrically processing information. Eil and Rao (2011) investigate how subjects update differently between neutral and ego-relevant information like beauty and intelligence. They find that subjects respond less when the information

doi:10.6342/NTU202002331

is bad (suggesting one's beauty or intelligence is inferior) and this effect only occurs in non-neutral settings. However, Coutts (2019) do not find the "good news-bad news effect" in their experiments. How people process ego relevant information is still debating.

Besides, our paper is an extension of rich literature of confirmation bias, which was documented in economics (Babcock et al., 1995) and psychology (Lord et al., 1979). Confirmation bias describes people's tendency to interpret the information in a fashion that is biased toward confirming one's prior belief. Glaeser and Sunstein (2013) introduces a model to show how balanced information leads to polarization. They suggest that the same information have diametrically opposite effect for those who have confirming and conflicting priors. Our experiment provides experimental evidence and illustrates a possible mechanism of this phenomenon.

doi:10.6342/NTU202002331

2 Experimental Design

There are ten rounds in the experiment, each round consists of two phases. At the very beginning of each round, the computer randomly decides the underlying states of two urns, urn A and urn B. The state space is two possible distributions that each urn could draw from, as described below. In the first phase, each subject receives a piece of information about the urn assigned to him/her. With this information, the subject is incentivized to truthfully report the probability of the true states for both urns. In the second phase, each subject observes another subject's elicited probability of their assigned urn (which is independently drawn to be the same or different from the original subject's urn). With this additional piece of information, each subject predicts the true state of both urns again.[1]

2.1 Design Details

In the first phase, subjects are independently assigned to either urn A or B with equal chance. Both urns contain one hundred digital balls, each ball is labeled by a number from 1 to 100. For each subject, the computer draws with replacement two balls randomly from the assigned urn. However, only one ball will be revealed to the subject. The ball number revealed to the subject is determined by one of two rules: The computer either reveals the larger one (*Maximum Rule*) or the smaller one (*Minimum Rule*). In each round, urn A and urn B are independently randomized to either follow the *Maximum Rule* or the *Minimum Rule* with equal chance. Subjects are not told the realization of the states (which rules the urns follow), but those

[1]Alternative experimental designs that were considered, but not implemented are listed in Appendix C.

doi:10.6342/NTU202002331

assigned to the same urn experience the same rule. After observing one ball, subjects have to predict the probability that the *Maximum Rule* is applied to urn A. Similarly, they also have to predict the probability that *Maximum Rule* is applied to urn B.

In the second phase, for each subject, the computer randomly chooses another subject, and reports his/her prediction of the urn they are assigned to. However, subjects do not know if this other subject was assigned to "Urn A" or "Urn B." In other words, each subject observes a number, which is the prediction of another subject in the first phase. After seeing the information from another subject, subjects again predict the probability that the *Maximum Rule* is applied to each urn.

2.2 Belief Elicitation

We use a two-stage menu of lottery choices as the belief elicitation mechanism in the experiment. Essentially, it is the Becker-DeGroot-Marschak (BDM) pricing procedure but easier for subjects to understand. In the first stage, subjects choose from a list of lottery pairs, which are choices between a random lottery and an event lottery. The random lotteries have winning probabilities ranging from 0%, 10%, ..., to 100%. Thus, subjects compare the probability of each random lottery with their beliefs that the event would occur. We allow only one "switching point" when completing the list of lottery pair. Based on the "switching point", subjects decide a second digit of probability in the second stage. After the decision is done, the computer randomly draws one number from 0 to 100. The lottery is chosen according to the drawn number, and the payoff is determined by the corresponded lottery.

doi:10.6342/NTU202002331

This method is incentive compatible. Without loss of generality, suppose one under-reports her beliefs from her real belief, 80%, to misreported belief, 60%. The results is the same when the drawn number is less than 60 and greater than 80. However, it is disadvantage for her if the number falls into the interval between 60 and 80. Since the event lottery will be chosen if she truthfully report the belief, and in that case the probability of getting the prize is 80%. Conversely, in the case of misreporting, the random lottery will be chosen and its probability of getting the prize is between 60% and 80%. Therefore, truthfully report the belief is the best interest for subjects.

Holt and Smith (2016) compared three mechanisms of belief elicitation and discussed the advantage of the two-stage menu of lottery choices. They find beliefs elicited from the two-stage menu to be more accurate and with lower average belief error in terms of Bayesian prediction.

2.3 Experimental Procedures

All sessions were conducted at Taiwan Social Sciences Experimental Laboratory (TASSEL), National Taiwan University (NTU). Six sessions were run during October 2019 and November 2019, for a total of 123 subjects. We recruited NTU students subjects using the TASSEL website powered by ORSEE (Greiner, 2015). Each session lasted approximately 100 minutes, and average earnings were 512 NT dollars (approx. $17). The experiment was programmed with z-Tree (Fischbacher, 2007) and conducted in Chinese. The experimental interfaces are shown in Figure 1a for the first stage and Figure 1b for the second stage of elicitation processes.

doi:10.6342/NTU202002331

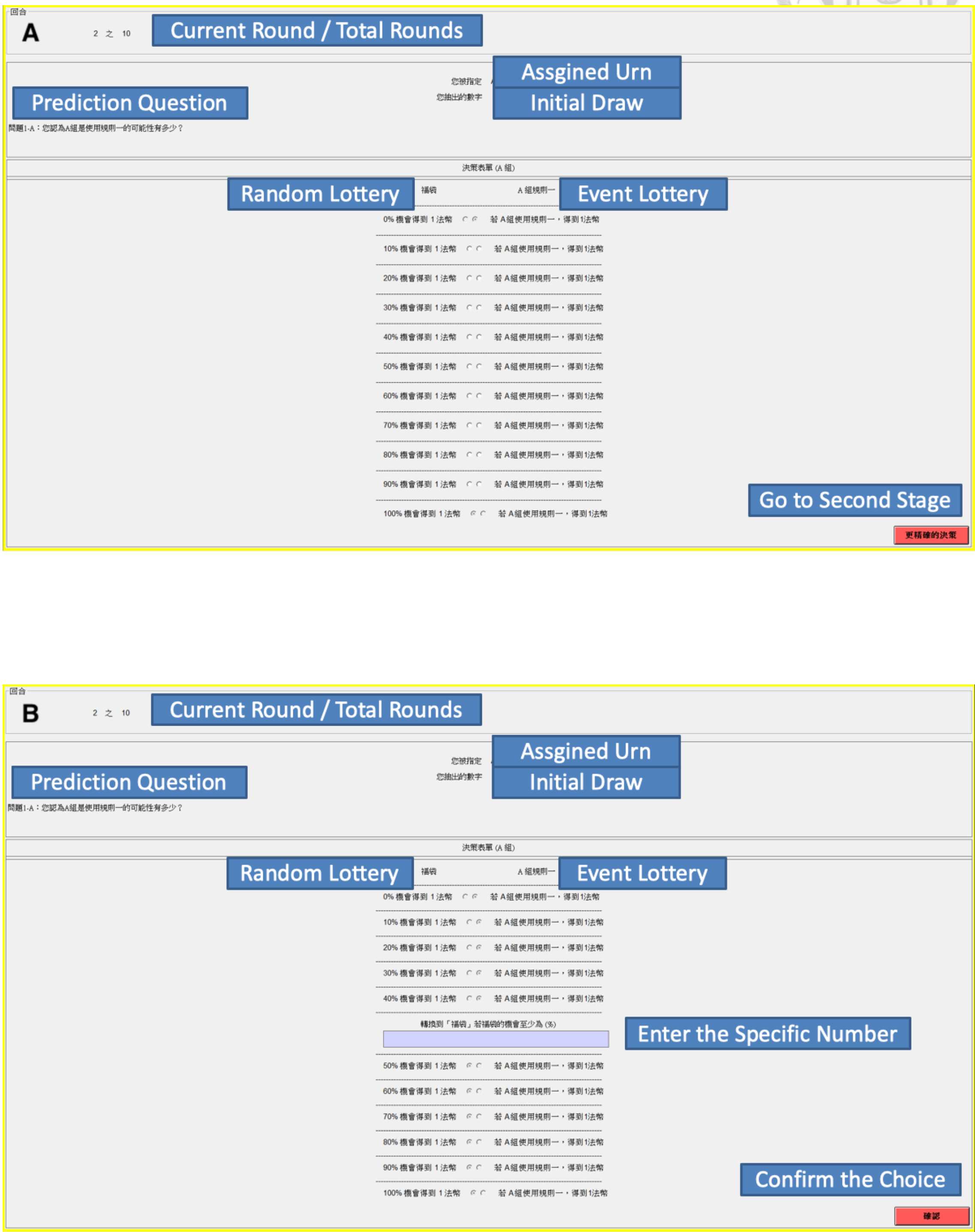


Figure 1: Two-stage Menu of Lottery Choices: (a) 1st Stage, and (b) 2nd stage.

doi:10.6342/NTU202002331

2.4 Bayesian Probability Predictions

For notation simplicity, we let urn A be the assigned urn and urn B be the irrelevant urn. We use $\theta_{\max}$ and $\theta_{\min}$ to denote the *Maximum Rule* and *Minimum Rule* of the assigned urn; the other urn also has two states, *Maximum Rule* and *Minimum Rule*, indicated by $\omega_{\max}$ and $\omega_{\min}$. The information s_1 denotes the observed ball in the first phase, s_2 is elicited probability of the assigned urn from another subject observed in the second phase.

2.4.1 The Structure of Two States

To calculate the Bayesian probability, we consider the structure of two possible states in advance. Consider the probability $\Pr(s_1|\theta_{\max})$ of seeing s_1 under *Maximum Rule* in the assigned urn. For two randomly drawn balls S_1^1 and S_1^2, there are two mutually exclusive events: Either the first drawn ball S_1^1 is the observed ball and therefore the second drawn ball is smaller than the observed ball, or exactly the opposite, that is, the second drawn ball S_1^2 is the observed ball and the first drawn ball is equal to or smaller than the observed ball. Therefore, the probability $\Pr(s_1|\theta_{\max})$ is:

$$\begin{aligned}\Pr(s_1|\theta_{\max}) &= \Pr\left(\left\{S_1^1 = s_1 \cap S_1^2 < s_1\right\} \vee \left\{S_1^1 \leq s_1 \cap S_1^2 = s_1\right\}\right) \\ &= \Pr(S_1^1 = s_1)\Pr(S_1^2 < s_1) + \Pr(S_1^1 \leq s_1)\Pr(S_1^2 = s_1) \\ &= \frac{1}{100}\cdot\frac{s_1 - 1}{100} + \frac{s_1}{100}\cdot\frac{1}{100} = \frac{2s_1 - 1}{10000} \end{aligned} \tag{1}$$

Similarly, the other probability is $\Pr(s_1|\theta_{\min}) = (201 - 2s_1)/10000$. Therefore, the probability distribution of observing the ball S_1 is linear under both the *Maximum*

doi:10.6342/NTU202002331

Rule (increasing by 0.02% from 0.01% when observing 1 to 1.99% when observing 100) and *Minimum Rule* (decreasing by 0.02% from 1.99% when observing 1 to 0.01% when observing 100).

2.4.2 Phase 1

In the first phase, the processed information is the observed ball, which is only useful to infer the state of urn A. With the observed ball, the Bayesian probability prediction for urn A is as follows.

$$\Pr(\theta_{\max}|s_1) = \frac{\Pr(s_1|\theta_{\max})\Pr(\theta_{\max})}{\Pr(s_1)} = \frac{(2s_1-1)/10000}{1/100}\cdot\frac{1}{2} = \frac{s_1}{100} - \frac{1}{200} \quad (2)$$

The Bayesian probability prediction for urn A shows that subjects should exactly predict at the percentage of their observed balls if they update the information by Bayesian Theorem. For example, suppose the observed ball s_1 is 30, the Bayesian probability is $\Pr(\theta_{\max}|s_1 = 30) = \frac{30}{100} - \frac{1}{200} = 29.5\%$.

The intuition of this prediction is simple. Given the observed ball, the probability of *Maximum Rule* for urn A is only depends on the other "unobserved ball". Thus, it is equivalent to the probability that the unobserved ball is smaller or equal to the observed ball, results in the term $s_1/100$. The subtraction of 1/200 (0.5%) is representing the tie case, in which both drawn balls are the same as the observed ball; therefore, it could also be the state of *Minimum Rule*. Thus, the 1% is split to both cases. The Bayesian probability prediction for urn B is straightforward since there is no information about urn B. As a results, $\Pr(\omega_{\max}|s_1)$ should be 0.5.

doi:10.6342/NTU202002331

2.4.3 Phase 2

In the second phase, subjects see another ball s_2, which is either from urn A or urn B. Because the actual source is unknown, it is useful to make inferences of both urns. Without loss of generality, we assume urn A is the assigned urn and urn B is the irrelevant urn, and their Bayesian probabilities in the second phase are:

$$\begin{aligned}\Pr(\theta_{\max}|s_1,s_2) &= \frac{\Pr(s_1\cap s_2|\theta_{\max})\cdot\Pr(\theta_{\max})}{\Pr(s_1\cap s_2)}\\ &= \frac{\Pr(s_2|s_1,\theta_{\max})\cdot\Pr(s_1|\theta_{\max})\cdot\Pr(\theta_{\max})}{\Pr(s_2|s_1,\theta_{\max})\cdot\Pr(s_1|\theta_{\max})\cdot\Pr(\theta_{\max})+\Pr(s_2|s_1,\theta_{\min})\cdot\Pr(s_1|\theta_{\min})\cdot\Pr(\theta_{\min})}\end{aligned} \tag{3}$$

$$\begin{aligned}\Pr(\omega_{\max}|s_1,s_2) &= \frac{\Pr(s_1\cap s_2|\omega_{\max})\cdot\Pr(\omega_{\max})}{\Pr(s_1\cap s_2)}\\ &= \frac{\Pr(s_2|s_1,\omega_{\max})\cdot\Pr(s_1|\omega_{\max})\cdot\Pr(\omega_{\max})}{\Pr(s_2|s_1,\omega_{\max})\cdot\Pr(s_1|\omega_{\max})\cdot\Pr(\omega_{\max})+\Pr(s_2|s_1,\omega_{\min})\cdot\Pr(s_1|\omega_{\min})\cdot\Pr(\omega_{\min})}\end{aligned} \tag{4}$$

where $\Pr(s_2|s_1,\theta_{\max})$

$$\begin{aligned}&= \Pr(s_2|s_1,\theta_{\max},\omega_{\max})\cdot\Pr(\omega_{\max}|s_1,\theta_{\max})+\Pr(s_2|s_1,\theta_{\max},\omega_{\min})\cdot\Pr(\omega_{\min}|s_1,\theta_{\max})\\ &= \Pr(s_2|s_1,\theta_{\max},\omega_{\max})\cdot\frac{1}{2}+\Pr(s_2|s_1,\theta_{\max},\omega_{\min},s_2\text{ from A})\cdot p_A\cdot\frac{1}{2}\\ &\qquad+\Pr(s_2|s_1,\theta_{\max},\omega_{\min},s_2\text{ from B})\cdot p_I\cdot\frac{1}{2}\end{aligned} \tag{5}$$

Thus, we have

$$\begin{aligned}\Pr(s_2|s_1,\theta_{\max}) &= \frac{2s_2-1}{10000}\cdot\frac{1}{2}+\frac{2s_2-1}{10000}\cdot\frac{1}{2}\cdot\frac{1}{2}+\frac{201-2s_2}{10000}\cdot\frac{1}{2}\cdot\frac{1}{2}\\ &= \frac{3}{4}\cdot\left(\frac{2s_2-1}{10000}\right)+\frac{1}{4}\cdot\left(\frac{201-2s_2}{10000}\right)\\ \Pr(s_2|s_1,\theta_{\min}) &= \frac{1}{4}\cdot\left(\frac{2s_2-1}{10000}\right)+\frac{3}{4}\cdot\left(\frac{201-2s_2}{10000}\right)\end{aligned} \tag{6}$$

doi:10.6342/NTU202002331

Equation 5 indicates the weightings that s_2 is under *Maximum Rule* or *Minimum Rule*. Since it is given the state of A is *Maximum Rule*, $\theta_{\max}$, only the state of B remains uncertain. By the settings of experimental design, there is equal chance that s_2 is either from urn A or urn B. It is the only possibility that s_2 is drawn under *Minimum Rule* when s_2 is from urn B and urn B is applied to *Minimum Rule*. Therefore, s_2 is drawn under *Maximum Rule* with 75% chance and *Minimum Rule* with 25% chance. With similar reason, we can also derive the probability in equation 6. The combination of probabilities (p_A, p_I) is the weights of the information source, indicating that the probability that new information is from the assigned urn or irrelevant urn. It is (0.5, 0.5) since the randomly drawn subject has equal chance to be assigned to urn A or B.

The following equations show the results of $\Pr(s_2|s_1,\omega_{\max})$ and $\Pr(s_2|s_1,\omega_{\min})$.

$$
\begin{aligned}
&\Pr(s_2|s_1,\omega_{\max})\\
&\quad= \Pr(s_2|s_1,\omega_{\max},\theta_{\max})\cdot\Pr(\theta_{\max}|s_1,\omega_{\max}) + \Pr(s_2|s_1,\omega_{\max},\theta_{\min})\cdot\Pr(\theta_{\min}|s_1,\omega_{\max})\\
&\quad= \Pr(s_2|s_1,\omega_{\max},\theta_{\max})\cdot\frac{2s_1-1}{200} + \Pr(s_2|s_1,\omega_{\max},\theta_{\min}, s_2 \text{ from A})\cdot p_A\cdot\frac{201-2s_1}{200}\\
&\qquad + \Pr(s_2|s_1,\omega_{\max},\theta_{\min}, s_2 \text{ from B})\cdot p_I\cdot\frac{201-2s_1}{200}\\
&s = \frac{2s_2-1}{10000}\cdot\frac{2s_1-1}{200} + \frac{201-2s_2}{10000}\cdot\frac{1}{2}\cdot\frac{201-2s_1}{200} + \frac{2s_2-1}{10000}\cdot\frac{1}{2}\cdot\frac{201-2s_1}{200}\\
&\quad= \frac{2s_1-1}{200}\cdot\left(\frac{2s_2-1}{10000}\right) + \frac{201-2s_1}{200}\cdot\left(\frac{1}{100}\right)
\end{aligned}
\tag{7}
$$

$$
\begin{aligned}
&\Pr(s_2|s_1,\omega_{\min})\\
&\quad= \frac{2s_1-1}{200}\cdot\left(\frac{1}{100}\right) + \frac{201-2s_1}{200}\cdot\left(\frac{201-2s_2}{10000}\right)
\end{aligned}
\tag{8}
$$

Equation 7 also shows the weightings that s_2 is under *Maximum Rule* or *Minimum*

doi:10.6342/NTU202002331

Rule but given the state of urn B, $\omega_{\max}$, instead of the state of urn A, $\theta_{\max}$. We can divide the equation into two parts, the state of urn A is either *Maximum Rule* or *Minimum Rule*. First of all, when the state of urn A is *Maximum Rule*, with the probability derived in equation 3, it is for sure that s_2 is drawn under *Maximum Rule*. Secondly, when the state of urn A is *Minimum Rule*, there is equal chance to draw s_2 under *Maximum Rule* or *Minimum Rule*. Thus, the probability of observing s_2 given states of u two urns $\omega_{\max}$ and $\theta_{\min}$ is the same as the probability of observing s_2, 1%. Equation 8 is derived by the same thoughts.

Hence, the Bayesian probability prediction for urn A (the assigned urn) is:

$$\begin{aligned}&\Pr(\theta_{\max}|s_1, s_2)\\&= \frac{[3(2s_2-1)+(201-2s_2)]\,(2s_1-1)}{[3(2s_2-1)+(201-2s_2)]\,(2s_1-1)+[(2s_2-1)+3(201-2s_2)]\,(201-2s_1)}\end{aligned} \tag{9}$$

By substituting equation 7 and 8 into 4, the Bayesian probability prediction for urn B (the irrelevant urn) is as follows.

$$\begin{aligned}&\Pr(\omega_{\max}|s_1, s_2)\\&= \frac{(2s_2-1)(2s_1-1)+100\cdot(201-2s_1)}{(2s_2-1)(2s_1-1)+100\cdot(201-2s_1)+100\cdot(2s_1-1)+(201-2s_2)(201-2s_1)}\end{aligned} \tag{10}$$

Alternatively, we can derive probabilities, $\Pr(s_2|s_1,\theta_{\max})$ and $\Pr(s_2|s_1,\omega_{\max})$ by the source of other's information. It is beneficial for analyzing how subjects consider other's information. Equation 11 and 12 show above concept. Exploiting the first

doi:10.6342/NTU202002331

ball and consequent beliefs, subjects form probabilities that second ball comes from urn A and B. Depends on two balls, subjects may distort p_A and p_I, both of which are 0.5 and p_A is equal to $(1-p_I)$ in theory.

$$\begin{aligned}
\Pr(s_2|s_1,\theta_{\max}) &= \Pr(s_2|s_1,\theta_{\max}, s_2 \text{ from A}) \cdot p_A + \Pr(s_2|s_1,\theta_{\max}, s_2 \text{ not from A}) \cdot (1-p_A) \\
&= \frac{2s_2-1}{10000} \cdot p_A + \frac{1}{100} \cdot (1-p_A) \qquad (11) \\
\Pr(s_2|s_1,\omega_{\max}) &= \Pr(s_2|s_1,\omega_{\max}, s_2 \text{ from I}) \cdot p_I + \Pr(s_2|s_1,\omega_{\max}, s_2 \text{ not from I}) \cdot (1-p_I) \\
&= \frac{2s_2-1}{10000} \cdot p_I + \frac{1}{100} \cdot (1-p_I) \qquad (12)
\end{aligned}$$

In equation 5 and 7, it is assumed that subjects update posteriors of two urns together. In other words, they rationally assign probabilities p_A and p_I so that the sum of p_A and p_I is always equal to 1. Thus, if the information is considered very unlikely being drawn from urn A, subject should put higher weight on urn B. Unfortunately, subjects may not be able to allocate probabilities p_A and p_I properly. For example, even if they believe the information has 10% chance coming from their assigned urn, they might only assign 60% to the irrelevant urn. One possible and intuitive updating process is that they separately update two urns. Specifically, when they deem the information not from one urn, they do not attribute it to the other urn. In fact, it is useless to subjects when updating the belief. In this situation, it seems that the information is drawn from an urn in which each ball is drawn with equal probability. In other words, when subjects regard the information is from the "useless urn", it provide no further clue for updating. To derive the theoretical prediction, the differences are caused by $\Pr(s_2|s_1,\theta_{\max})$, $\Pr(s_2|s_1,\theta_{\min})$,

doi:10.6342/NTU202002331

$\Pr(s_2|s_1, \omega_{\max})$, and $\Pr(s_2|s_1, \omega_{\min})$. Therefore, the theoretical results are as follows.e

$$\begin{aligned}
&\Pr(\theta_{\max}|s_1, s_2) \\
&= \frac{[(2s_2-1)p_A + 100(1-p_A)](2s_1-1)}{[(2s_2-1)p_A + 100(1-p_A)](2s_1-1) + [(201-2s_2)p_A + 100(1-p_A)](201-2s_1)}
\end{aligned} \tag{13}$$

$$\begin{aligned}
&\Pr(\omega_{\max}|s_1, s_2) \\
&= \frac{[(2s_2-1)p_I + 100(1-p_I)](2s_1-1)}{[(2s_2-1)p_I + 100(1-p_I)](2s_1-1) + [(201-2s_2)p_I + 100(1-p_I)](201-2s_1)}
\end{aligned} \tag{14}$$

doi:10.6342/NTU202002331

3 Results

3.1 Adherence to Bayesian Updating

3.1.1 Compliance After Initial Draw

Figure 2a presents elicited probabilities of the assigned urn after drawing a ball in the first phase. Each data point represents the reported belief of a subject in a particular round. The majority of data are very close to the correct Bayesian posteriors, with nearly 90 percent of the data aligned with the theory if we allow for an errors margin of plus and minus 10 percentage points (±10%).[2] The elicited probabilities of the irrelevant urn, in which they do not have any information, are shown on Figure 2b, in which over 80% of the elicited probabilities are between 0.4 and 0.6 (50% ± 10%). Table 1 shows that a majority of choices conform with the theoretical predictions as we reduce the margin of error allowed. Even under the strictest case allowing for only 1 percentage point error (±1%), 60% and 55% of the choices are considered Bayesian in the assigned and the irrelevant urn, respectively.

The squares in Figure 2 represent the mean elicited probabilities averaged across all subjects with the same initial draw. They closely adhere to the Bayesian posteriors, especially for the assigned urn. Notice that there is a cluster of elicited probabilities along the 45-degree line in Figure 2b, implying that some subjects also use the initial draw to update the irrelevant urn. We find that those choices come from one-time behavior of different subjects and not concentrated in partic-

[2] Alternatively, one could construct the upper and lower bounds relative to the initial draw. For example, allowing for a 10 percent error results in 50% ± 5% for the ball 50, but 10% ± 1% for the ball 10. This criteria is harsh to those who draw a very small or large ball since they have stronger information. However, under it 76% of the data are still considered to be aligned with theory.

doi:10.6342/NTU202002331

ular rounds, indicating that they are not caused by particular subjects or rounds.[3] Although these choices consists of only 3% of the data, they inflate the correlation between the elicited probabilities of the assigned and irrelevant urn.[4] Without these choices, the correlation is 0.003 ($p > 0.1$), indicating that the vast majority of probabilities are elicited with the knowledge that states of the two urns are independent.[5] In conclusion, most of the choices are consistent with Bayesian updating derived in section 2.4.2.

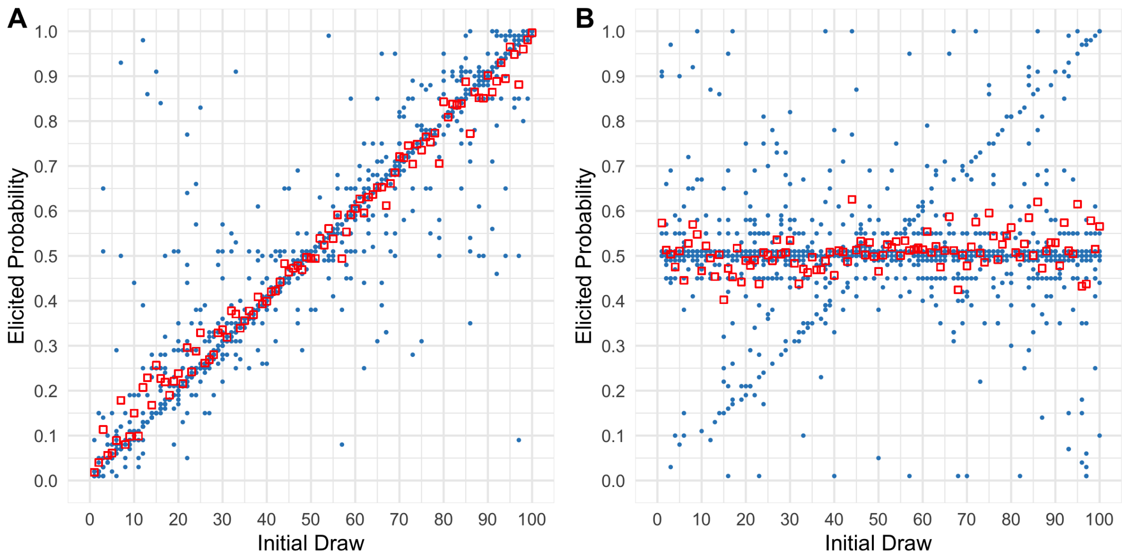


Figure 2: Elicited Beliefs in the First Phase of the (a) Assigned (b) Irrelevant Urn

Table 1: Percentage of Theory-consistent Choices Under Different Error Margins

Error Margin	Assigned Urn	Irrelevant Urn
±10 percentage points	89%	81%
±5 percentage points	81%	74%
±3 percentage points	75%	60%
±1 percentage points	66%	55%

[3] See Appendix A for further details.

[4] A total of 37 choices lie exactly on the 45-degree line excluding initial draws between 40 and 60 where we cannot easily tell if they updated beliefs of the irrelevant urn or not.

[5] Similarly, the second phase correlation between the two urns is 0.006 ($p > 0.1$). Computing with all data, the first and second phase correlations are 0.067 and 0.029, respectively.

doi:10.6342/NTU202002331

3.1.2 Failure After Observing New Information

There exists one intuitive difference between the two possible states of the urn: When the true state is the *Maximum Rule*, the subject is more likely to observe a ball larger than 50, while under the *Minimum Rule*, the subject is more likely to observe a ball equal to or smaller than 50. This leads to a straightforward heuristic for subjects to determine whether new information in the second phase is more likely to come from an urn under the *Maximum Rule* or *Minimum Rule*. As a result, we classify the second-phase information coming from another subject, as either *confirming* or *conflicting* information. In particular, the new information is *confirming* if first and second phase information are both within 1–50 or both within 51–100, while it is *conflicting* when one is within 1–50 and the other one is within 51-100.[6]

Compared to the first phase, belief-updating in the second phase is much worse.[7] Figure 3 summarizes the distribution of Bayesian posteriors and the average deviation from them on different intervals. When the new information is *confirming*, we find that subjects deviate less in the assigned urn, but deviate more in the irrelevant urn. This suggests that it is easier to correctly process new information regarding the assigned urn that aligns with what subjects already have. In contrast, updating behavior for the irrelevant urn is far from the Bayesian prediction as the overall deviations are larger than the assigned urn (Figure 3b).

Furthermore, the R-squared predicting elicited probabilities using Bayesian posteriors shows that subjects perform updating well in the assigned urn when the

[6]Some information may be too close to 50 to be "confirming" or "conflicting" enough, such as initial draws or new information between 40 and 60. Excluding these cases, we expect to find stronger effects.

[7]See Figure 10 of Appendix B for the raw data plotted like Figure 2.

doi:10.6342/NTU202002331

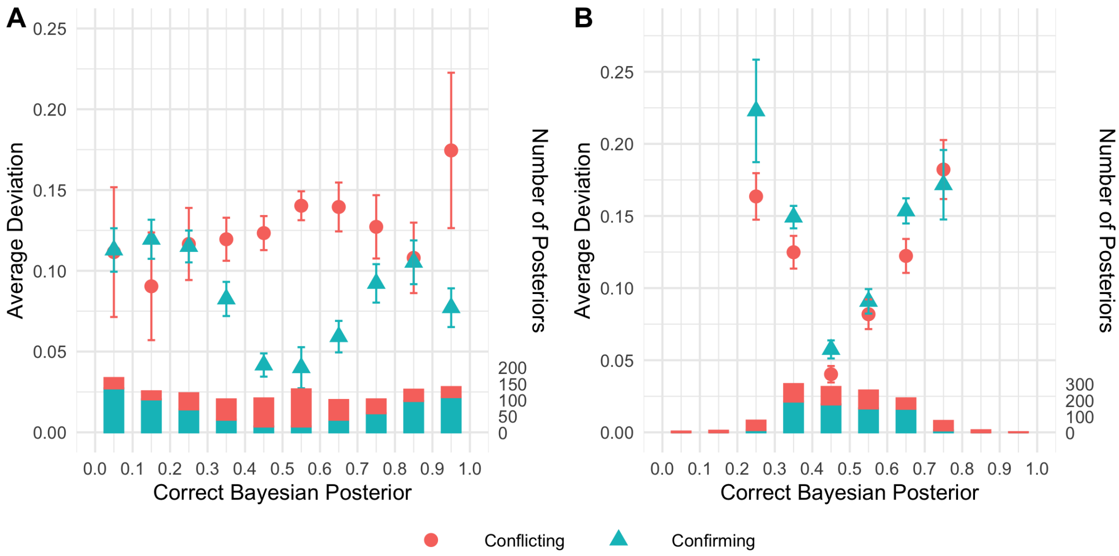


Figure 3: Elicited Beliefs Distribution in the Second Phase of (a) the Assigned, and (b) Irrelevant Urn

information is *confirming* ($R^2 = 0.82$), but perform worse when it is *conflicting* ($R^2 = 0.51$). In contrast, for the irrelevant urn, subjects perform worse when the new information is confirming ($R^2 = 0.33$), but perform better when it is conflicting ($R^2 = 0.52$). The differences in R^2 are statistically significant for both urns (variance ratio test, $p < 0.001$). The results in Appendix B show that the slopes between *confirming* and *conflicting* information are not significantly different in Figure 10a ($p = 0.175$) and Figure 10b ($p = 0.434$).[8]

3.2 The Echo Chamber

In principle, subjects should update their beliefs of both urns regardless of the information received in the second phase because there is always a chance the new information could be from either urn. However, the irrelevant urn has the natural

[8] We test the coefficient β_3 from the model: $Beliefs\beta_0 + \beta_1 Bayesian + \beta_2 Confirming + \beta_3 Interaction + \epsilon$, where the dummy variable *Confirming* indicates the new information is confirming (=1) or not (=0), *Interaction* is the interaction term of *Bayesian* and *Confirming*.

doi:10.6342/NTU202002331

advantage that one should only update it according to the new information regarding the ball of the second phase, since the first ball only carries information about the assigned urn. Therefore, we can easily infer how subjects attribute new information to each urn in the second phase from their updating behavior.

Figure 4 plots elicited probabilities against second-phase information.[9] The red dots are elicited beliefs around 0.5, adhering to the Bayesian prediction of the first phase, indicating "fully dissociate" subjects who do not update irrelevant urn beliefs at all (and should completely attribute the new information to the assigned urn). On the other hand, the blue crosses along the 45-degree line indicate "fully attribute" types who completely ignore the fact that there is some probability that the new information is from their assigned urn.[10] These two types are strongly biased since they put extreme weight on the new information when updating the irrelevant urn. However, they account for 76.7% of the choices when we allow 5 percentage points of error. The intermediate types with more reasonable weights are shown as green triangles in Figure 4, but consist only 18.7% of the choices. This includes those who follow Bayesian updating. Lastly, the remaining 4.6% of choices in black are difficult to rationalize, and might reflect confusion or some other information processing method. We summarize the updating behavior in the Table 2.

In Figure 5, we separate second-phase information into *confirming* and *conflicting* information as defined in section 3.1.2. To compare the difference in behavior between receiving confirming and conflicting information, we use a dummy indi-

[9]We drop the choices if their first phase beliefs of the irrelevant urn are out of the range, $[0.45, 0.55]$. The remaining choices plotted in the Figure 4 contain 74% of the data.

[10]The purple dot-cross symbols are overlapping area of the two types, in which we cannot distinguish their types.

doi:10.6342/NTU202002331

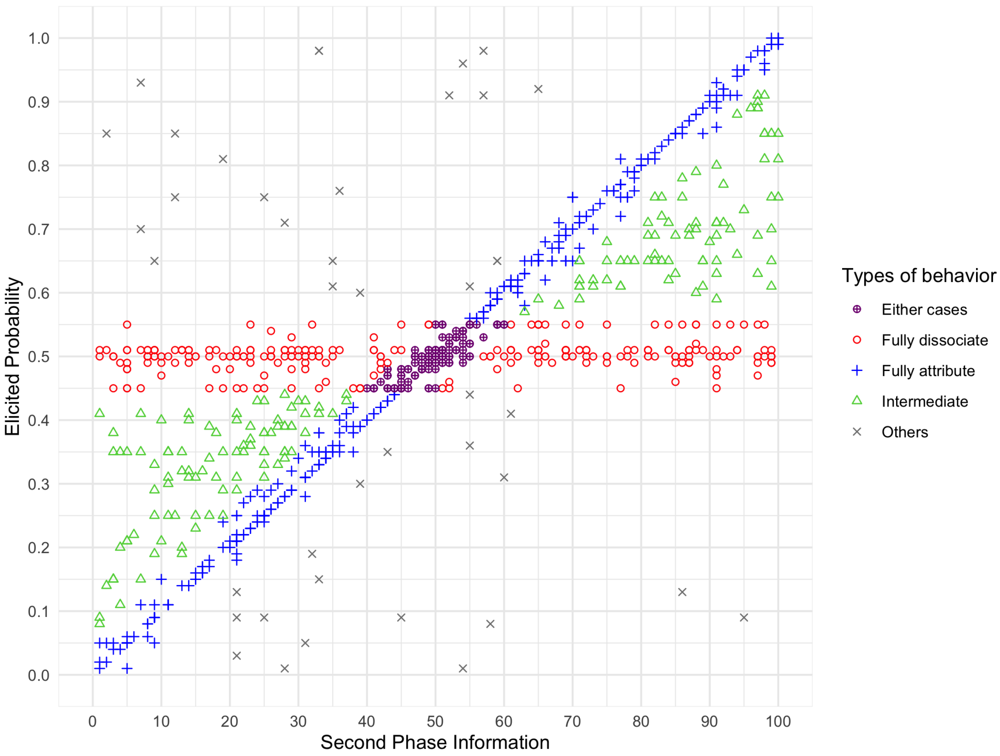


Figure 4: Types of Behavior (Irrelevant Urn)

Table 2: Types of Behavior (Irrelevant Urn)

Types of Choices	Definition	Percentage
Either	Either fully dissociate or fully attribute type.	16.3 %
Fully Dissociate	Other subject's information comes from the assigned urn.	25.4 %
Fully Attribute	Other subject's information comes from the irrelevant urn.	35 %
Intermediate	Put reasonable weights on other subject's information	18.7 %
Others	Choices cannot be classified into above four types.	4.6 %

doi:10.6342/NTU202002331

cating confirming information to predict the occurrence of two distinct types of behavior, completely attribute the information to the assigned urn (Fully Dissociate) and the irrelevant urn (Fully Attribute). Table 3 report fixed-effect panel regression results clustered at the subject level, predicting whether the inferred prior belief fully attributes the new information to the irrelevant urn using whether information is *confirming* or not. For *confirming* information, 33.7% of the choices completely attribute the new information to the assigned urn, while 31.1% of the choices completely attribute the new information to the irrelevant urn. However, when subjects receive *conflicting* information, only 16.5% of the choices attribute new information to the assigned urn, significantly lower than that under *confirming* information. Moreover, 39% of the choices completely attribute new information to the irrelevant urn, significantly higher than that under *confirming* information. This results demonstrates a confirmation bias where subjects overweight (underweight) the possibility that new information came from the assigned urn when it confirms (refutes) their prior.

Table 3: Attribution of the Information

Fully Attribute to	(1) Assigned Urn	(2) Irrelevant Urn
Confirming Information	0.165***	-0.079***
	(0.022)	(0.025)
Constant	0.172***	0.390***
	(0.017)	(0.019)
N	914	914

Note: Standard errors in parentheses, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Among those who completely attribute the new information to the irrelevant urn (Fully Attribute), their updated beliefs of the assigned urn should remain unchanged

doi:10.6342/NTU202002331

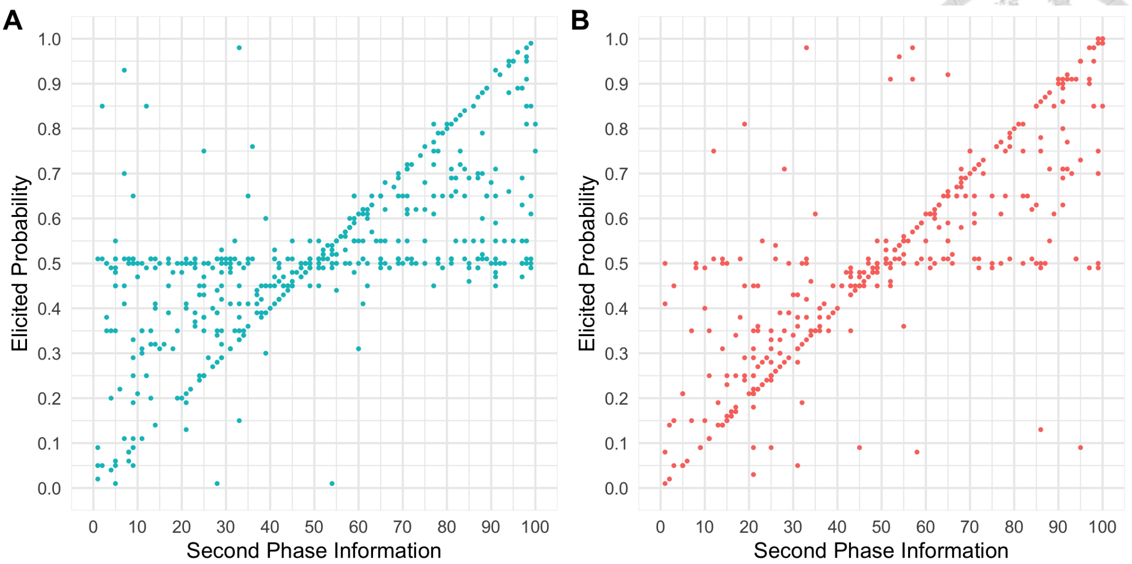


Figure 5: Elicited Beliefs of the Irrelevant Urn: (a) Confirming, and (b) Conflicting Information.

because they believe the information is coming solely from the irrelevant urn. Indeed, the posteriors of the assigned urn show that 75% do not update the assigned urn beliefs much.[11] The remaining 25% also changes their beliefs regarding the assigned urn, overreacting the new information.

In contrast, among those who completely attribute the new information to the assigned urn (Fully Dissociate), beliefs of the assigned urn should be updated as if they have two balls from that urn, resulting in a Bayesian updating process similar to equation (3) in section 2.3.2.[12] Unexpectedly, 54% of these choices stick to their first-phase posteriors of the assigned urn. This implies at least $25.4\% \times 54\% = 13.7\%$ of all choices completely ignore the new information and update neither urn.[13] Figure 6 plots the remaining choices after excluding those which completely ignore the

[11]This number is calculated by allowing 5% error. In fact, 63% have the exact same first and second posterior beliefs.

[12]The Bayesian prediction of having two balls from the same urn is: $\Pr(\theta_{\max}|s_1, s_2) = \Pr(s_2|\theta_{\max}) \cdot \Pr(\theta_{\max}|s_1) / [\Pr(s_2|\theta_{\max}) \cdot \Pr(\theta_{\max}|s_1) + \Pr(s_2|\theta_{\min}) \cdot \Pr(\theta_{\min}|s_1))]$.

[13]13.7% is the lower bound since 25.4% excludes choices when second phase information are close to 50 that could be either Fully Dissociate or Fully Attribute.

doi:10.6342/NTU202002331

new information. Figure 6a compares the elicited probabilities of fully dissociate types and the Bayesian posterior assuming that both balls came from the same urn. Even though subjects fully dissociate the information from the irrelevant urn, the updating behavior systematically under-weights the new information from the other subject, resulting in a slope of 0.67 that is significantly lower than 1 ($p < 0.001$). In fact, the elicited probabilities are closer to the Bayesian probability prediction derived in section 2.3.2 (Figure 6b), although the slope (0.78) is still lower than 1 ($p < 0.001$).

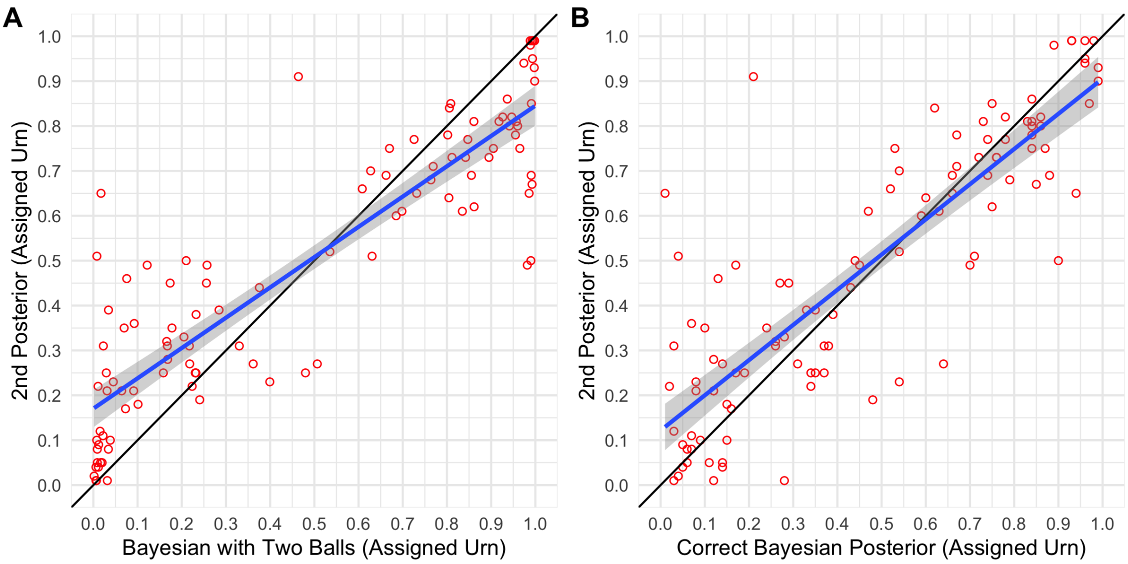


Figure 6: Fully Dissociate: (a) Two Balls from Assigned Urn. (b) Correct Bayesian.

doi:10.6342/NTU202002331

3.3 Inferred Prior Beliefs of Other's Information

In this section, we estimate the source beliefs (p_A, p_I), probabilities subjects consider the information comes from, which reflects how subjects attribute the information to the assigned and irrelevant urn. In our experiment, it is explicitly stated that the combination of source beliefs is (0.5, 0.5). We use the four posteriors elicited (first/second phase in the assigned/irrelevant urn) to estimate subjects' (p_A, p_I) by conducting a maximum likelihood estimation.[14] We follow a structural estimation method similar to that in Costa-Gomes and Crawford (2006) but impose a logit error structure instead of spike-logit because it is hard for subjects to exactly hit the Bayesian updating prediction given the complicated Bayesian calculation.

We allow for 21 possible types, ranging from $p_A = 0$, 0.05..., to 1.[15] We assume that each subject's updating behavior is fixed across the 10 rounds. Formally, let $k = 0, 5, ..., 100$ (which stands for the source belief p_A from 0%, 5%, ..., to 100%) index our types, $R = 20$ denote the total number of elicited probabilities (since each round consists of two updating decisions),[16] and x_r^i denote subject i's posteriors in choice r. Given subject's type and information received, let $t_r^{i,k}$ denote the predicted posterior for a type-k subject i in round r. In order to interpret the pattern of

[14]To properly investigate individual "updating" types, we use subjects' first posteriors to calculate the target second posteriors, otherwise it could be problematic for those who deviate from the Bayesian posteriors in the first phase. For example, subject who report 60% as posteriors of the irrelevant urn and 38% as posteriors of the assigned urn in both phases is actually behaving as an "ignoring" type in the second phase. However, if we use the correct Bayesian posteriors in the first phase as benchmarks to calculate the second phase posteriors, we will mistakenly believe this subject is perfectly Bayesian.

[15]It is unnecessary to divide the types further since different p_A would map into the same combination of balls. For example, suppose one subject has the balls 30 and 70 in the first and second phase, respectively. The Bayesian posteriors are 0.38 for the assigned urn and 0.61 for the irrelevant urn if $p_A = 0.5$. If $p_A = 0.51$, the corresponding posteriors hardly change, so we cannot distinguish the subject's type.

[16]We assume that all posteriors are updated independently.

doi:10.6342/NTU202002331

deviations from one's updating, we specify a logit error structure in which, in every particular round, a subject updates to the exact predicted posterior of one's type with highest probability, and the probability decreases as we move away from the predicted posterior. In particular, a type-k subject's assigned urn posterior in round r satisfies the logit density function $d_r^k(x_r^i, t_r^{i,k}, \lambda)$ with precision parameter λ:

$$d_r^k(x_r^i, t_r^{i,k}, \lambda) \equiv \frac{\exp\left[\lambda E(x_r^i | t_r^{i,k})\right]}{\sum_{z_r^i} \exp\left[\lambda E(z_r^i | t_r^{i,k})\right]}. \tag{15}$$

where the expected payoff $E(x|t_r^{i,k}) = x \cdot t_r^{i,k} + (1-x) \cdot (x+1)/2$, the actual payoff subjects earn in the experiment. Therefore, the density of a type-k subject with updates $\boldsymbol{x^i} \equiv (x_1^i, ..., x_R^i)$ is

$$d^k(x^i, t^{i,k}, \lambda) \equiv \prod_r d_r^k(x_r^i, t_r^{i,k}, \lambda). \tag{16}$$

Let p^k denote a subject's prior probability of being type-k, with $\sum_{k=1}^K p^k = 1$ and $\boldsymbol{p} \equiv (p^1, ..., p^K)$. By multiplying the right hand-side of (15) by p^k, summing over k and taking logarithms, the log-likelihood function for subject i becomes

$$\ln L(p, \varepsilon, s | x^i) = \ln \left[\sum_{k=1}^K p^k d^k(x^i, t^{i,k}, \lambda) \right]. \tag{17}$$

Given the estimate of λ, it is clear from (17) that the maximum likelihood estimate of p sets $p^k = 1$ for the generically unique k that yields the highest $d^k(x^i, t^{i,k}, \lambda)$. The maximum likelihood estimate of λ is the logistic scale parameter describing the spreading of subject's updating.

doi:10.6342/NTU202002331

Figure 7a shows that on average subjects assign different weights when facing conflicting and confirming information. The weight is $p_A = 32\%$ (median = 20%) when estimated using only rounds in which the information is conflicting, but it increases to $p_A = 44\%$ (median = 45%) when using rounds in which information in confirming. The difference of subject beliefs between confirming and conflicting is significant ($44\% \gg 32\%$: t-test $p < 0.001$; Wilcoxon signed-rank test $p = 0.003$), suggesting the occurrence of an echo chamber effect.

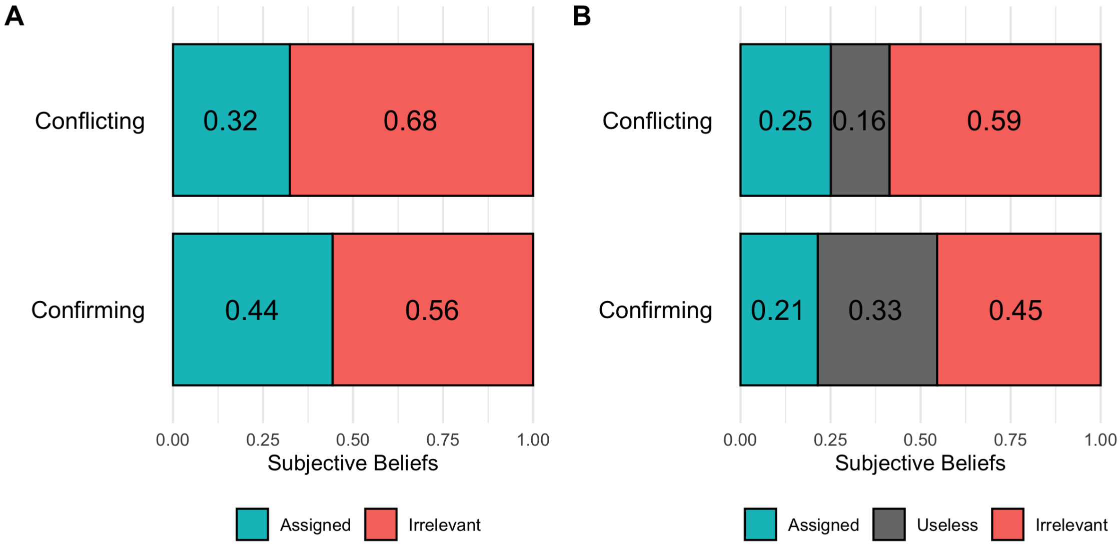


Figure 7: Models of Information Sources: (a) Two Urns (b) Three Urns.

The above model restricts the sum of p_A and p_I to necessarily equal to one, which implies the information must originate from either the assigned or irrelevant urn. This assumption adheres to our experimental design. However, people may underweight others' information. Also, notice that subjects do not always update correctly compared to Figure 2a. Therefore, subjects may believe that the information received does not coincide with a ball drawn from one of the two urns. As a result, they might decide to discount or even ignore this information completely

doi:10.6342/NTU202002331

when updating their beliefs in the second phase.

We can modify our model to accommodate the possibility of under-weighting information. Subjects may view the information as useless for making any inference, and thus ignore and attribute it to a "useless urn" added to our model to deal with such situations. If the information comes from the useless urn, each ball is drawn with equal probability. In other words, this information is completely random and not helpful to update any posteriors at all. The theoretical predictions of $\Pr(s_2|s_1, \theta_{\max})$ derived in equation (5) becomes[17]

$$\begin{aligned}
&\Pr(s_2|s_1, \theta_{\max}) \\
&\quad = \Pr(s_2|s_1, \theta_{\max}, \omega_{\max}) \cdot \Pr(\omega_{\max}|s_1, \theta_{\max}) + \Pr(s_2|s_1, \theta_{\max}, \omega_{\min}) \cdot \Pr(\omega_{\min}|s_1, \theta_{\max}) \\
&\quad = \frac{1}{2}\Big[\Pr(s_2|s_1, \theta_{\max}, \omega_{\max}, \text{Assigned } s_2) \cdot p_A + \Pr(s_2|s_1, \theta_{\max}, \omega_{\max}, \text{Irrelevant } s_2) \cdot p_I \\
&\qquad + \Pr(s_2|s_1, \theta_{\max}, \omega_{\max}, \text{Useless } s_2) \cdot p_U + \Pr(s_2|s_1, \theta_{\max}, \omega_{\min}, \text{Assigned } s_2) \cdot p_A \\
&\qquad + \Pr(s_2|s_1, \theta_{\max}, \omega_{\min}, \text{Irrelevant } s_2) \cdot p_I + \Pr(s_2|s_1, \theta_{\max}, \omega_{\min}, \text{Useless } s_2) \cdot p_U\Big].
\end{aligned} \tag{18}$$

Figure 7b shows that subjects are still significantly prone to attributing information to the irrelevant urn when it is conflicting (59% $\gg$ 45%: t-test: $p = 0.001$; Wilcoxon signed-rank test: $p = 0.002$). However, this effect disappears for the assigned urn—subject beliefs of the information source are not significantly different between conflicting and confirming information (25% $\sim$ 21%: t-test and Wilcoxon signed-rank test: $p > 0.1$). Instead, the effect is entirely on the useless urn, showing

[17]Equation 18 demonstrates how to break down the probability $\Pr(s_2|s_1, \theta_{\max})$ to three urns. We can also apply the same method to the remaining three required probabilities, $\Pr(s_2|s_1, \theta_{\min})$, $\Pr(s_2|s_1, \omega_{\max})$, and $\Pr(s_2|s_1, \omega_{\min})$.

doi:10.6342/NTU202002331

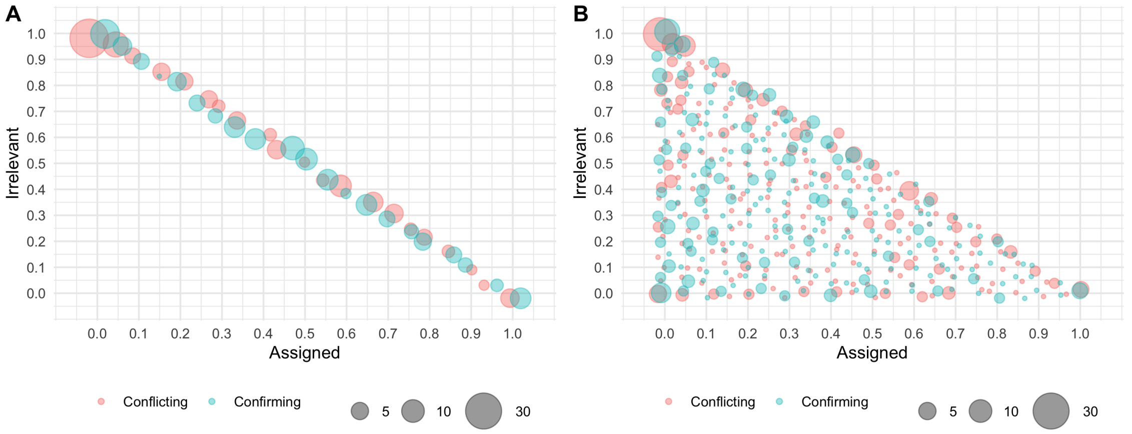


Figure 8: Information Sources Distributions: (a) Two Urns (b) Three Urns.

that subjects tend to ignore the information when it is confirming ($33\% \gg 16\%$: t-test: $p < 0.001$; Wilcoxon signed-rank test: $p < 0.001$). The distributions of subjects in the two models are shown in Figure 8, and individual beliefs of the source are listed in Table 6.

To illustrate the differential processing of confirming and conflicting information, we consider three representative types: Subjects who attribute the information completely to the assigned urn ($p_A = 1$), completely to the irrelevant urn ($p_A = 0$), and those close to Bayesian ($p_A = 0.5$). Applying the same maximum likelihood estimation with these 3 types ($p_A = 0, 0.5, 1$) instead of 21 types ($p_A = 0, 0.05, \cdots, 1$), we estimate individual types and classify subjects accordingly. The results shown in Table 4 indicate that 24.4% more subjects attribute the information completely to the assigned urn when it is confirming. In contrast, 10.6% more subjects attribute the information completely to the irrelevant urn when it is conflicting. Table 4 uncovers this alternation at the individual level. Subjects along the diagonal (49.6%) are consistent under both information. Importantly, the upper triangle subjects (37.4%, underlined) put more weight on the assigned urn when moving to confirming infor-

doi:10.6342/NTU202002331

mation (from conflicting information). In other words, these subjects exhibit an "echo chamber effect," since they are more likely to believe that confirming information comes from their assigned urn and vise versa.

Table 4: Individual Type Transition: Conflicting vs. Confirming (%)

		Confirming p_A			
Conflicting	p_A	0	0.5	1	Total
	0	21.1	21.1	12.2	54.4
	0.5	6.5	25.2	4.1	35.8
	1	1.6	4.9	3.3	9.8
Total		29.3	51.2	19.5	100.0

It is apparent that subjects are not necessary consistent between belief-updating of the assigned urn and the irrelevant urn. This may be caused by the inability to properly assign probabilities between the two urns. In particular, subjects could update the two urns independently, instead of comprehensively evaluate the information and simultaneously update their beliefs about the assigned and irrelevant urn. Hence, they utilize the information and assess the probability for it to come from each urn separately. If they deem the information irrelevant, it is attributed to a useless urn, in which each ball (1 to 100) is drawn with equal chance, instead of the other urn. Therefore, subjects assign underlying beliefs (p_A, p_U) and (p_I, p_U) when assessing the assigned and irrelevant urn, respectively.

We compare underlying beliefs p_A and p_I when receiving confirming and conflicting information. Specifically, we predict underlying beliefs with a constant and the dummy for *Confirming* information to predict p_A in each round, and cluster standard errors at the subject level to control for repeated observations. We exclude choices which could only be rationalized with impossible beliefs that are not in the interval $[0, 1]$, which happens more often for the irrelevant urn. This leaves us with

doi:10.6342/NTU202002331

846 observations for the assigned urn, in contrast to 775 observations for the irrelevant urn. Table 5 column (1) and (2) show that the directions of coefficients confirm the asymmetric updating. When the information is aligned with their priors, subjects put insignificantly more weight (2.4%) on the assigned urn, but significantly less (-18.7%, $p < 0.001$) weight on the irrelevant urn. However, notice that some information are more confirming or conflicting than others. For instance, when information is 51, one can hardly infers anything. Similarly, the information may not really be confirming or conflicting for subjects where the initial draws are close to 50. Thus, we regard information as strongly confirming or conflicting when neither the initial draw nor the new information are between 40 and 60. The results shown in column (3) and (4) indicated that the effects are even larger at 5.6% ($p < 0.05$) and -27.3% ($p < 0.001$) for the assigned and irrelevant urn, respectively.

Table 5: Independent Source Beliefs

Source Beliefs:	(1) Assigned Urn	(2) Irrelevant Urn	(3) Assigned Urn	(4) Irrelevant Urn
Confirming Information	0.024 (0.020)	-0.187*** (0.031)	0.056* (0.025)	-0.273*** (0.037)
Constant	0.155*** (0.019)	0.533*** (0.028)	0.142*** (0.022)	0.611*** (0.033)
Stronger Confirming/Conflicting	✗	✗	✓	✓
N	846	775	555	518

doi:10.6342/NTU202002331

Table 6: Individual Source Beliefs

	Two Urns		Three Urns					Two Urns		Three Urns			
	Conflicting	Confirming	Conflicting		Confirming			Conflicting	Confirming	Conflicting		Confirming	
ID	p_A	p_A	p_A	p_I	p_A	p_I	ID	p_A	p_A	p_A	p_I	p_A	p_I
416	0	0	0	0.25	0	1	621	0.25	0.2	0.25	0.75	0.2	0.8
111	0	0	0	1	0	1	620	0.25	0.45	0.05	0.8	0.1	0.4
115	0	0	0	1	0	1	512	0.25	0.5	0.05	0.55	0.05	0.2
508	0	0	0	1	0	1	307	0.25	0.5	0.15	0.6	0.45	0.45
519	0	0	0	1	0	1	417	0.25	0.65	0.25	0.75	0.1	0.25
604	0	0	0	1	0	1	404	0.3	0	0.3	0.6	0	1
616	0	0	0	1	0	1	109	0.3	0.45	0.3	0.7	0	0.25
212	0	0.05	0	1	0.05	0.65	503	0.35	0.35	0	0.8	0	0.25
504	0	0.05	0	1	0.05	0.95	221	0.35	0.5	0.35	0.65	0.25	0.45
511	0	0.05	0	1	0.05	0.95	407	0.35	0.55	0	0	0	0.05
217	0	0.1	0	0.95	0.05	0.85	502	0.35	0.6	0	0.4	0	0.25
301	0	0.1	0	1	0	0.85	613	0.35	0.9	0.4	0.55	0.5	0.05
313	0	0.1	0	1	0.1	0.9	316	0.4	0.4	0.4	0.6	0.4	0.6
607	0	0.2	0	0.95	0	0.55	213	0.4	0.55	0.3	0.6	0.45	0.35
210	0	0.2	0	0.95	0.1	0.8	509	0.45	0.1	0.3	0.55	0	0.65
619	0	0.2	0	1	0.2	0.75	601	0.45	0.4	0.4	0.45	0.3	0.55
218	0	0.2	0	1	0.2	0.8	317	0.45	0.45	0.45	0.55	0.2	0.45
412	0	0.35	0	0.95	0.35	0.65	617	0.45	0.5	0.45	0.55	0.4	0.35
108	0	0.35	0	1	0.05	0.15	610	0.45	0.65	0.45	0.55	0.65	0.35
614	0	0.35	0	1	0.1	0.45	517	0.45	0.75	0.45	0.55	0	0.1
611	0	0.35	0	1	0.2	0.65	117	0.5	0.55	0.5	0.5	0.3	0.25
310	0	0.4	0	0.9	0.35	0.6	214	0.55	0.5	0.5	0.45	0.1	0.2
516	0	0.4	0	1	0.15	0.45	211	0.55	0.7	0.1	0	0.05	0
320	0	0.45	0	0.45	0.1	0.4	622	0.6	0	0	0	0	1
202	0	0.45	0	1	0.3	0.5	311	0.6	0.3	0.6	0.4	0.3	0.7
314	0	0.65	0	0.85	0.35	0.2	312	0.6	0.35	0.6	0.4	0.35	0.65
103	0	0.65	0	0.95	0.6	0.25	521	0.6	0.4	0.6	0.4	0.2	0.55
414	0	0.65	0	1	0.05	0.25	319	0.6	0.45	0.6	0.4	0.45	0.55
102	0	0.7	0	1	0.65	0.2	507	0.6	0.45	0.6	0.4	0.45	0.55
624	0	0.75	0	0.5	0	0	513	0.6	0.45	0.6	0.4	0.45	0.55
306	0	0.8	0	1	0.3	0.1	625	0.6	0.55	0.05	0	0	0.2
501	0	0.85	0	1	0.5	0	603	0.6	0.55	0.35	0	0	0
208	0	1	0	1	0.4	0	118	0.65	0	0.65	0.35	0	1
203	0	1	0	1	0.5	0	114	0.65	0.35	0.55	0.25	0	0.4
216	0	1	0	1	1	0	406	0.65	0.45	0.4	0	0	0.1
205	0.05	0	0.05	0.95	0	0.8	201	0.65	0.45	0.5	0.25	0	0.35
318	0.05	0	0.05	0.95	0	1	411	0.65	0.5	0.55	0.3	0.15	0.35
615	0.05	0.25	0.05	0.95	0.05	0.65	104	0.65	0.65	0.65	0.35	0.4	0.35
321	0.05	0.3	0.05	0.95	0	0.5	403	0.65	1	0	0	0.8	0
116	0.05	0.5	0.05	0	0.3	0.5	520	0.7	0	0.2	0	0	1
606	0.05	0.5	0.05	0.95	0.5	0.5	608	0.7	0	0.7	0.3	0	1
515	0.05	0.55	0	0.95	0.1	0.35	612	0.7	0.4	0.7	0	0.4	0.6
609	0.05	0.65	0	0.95	0.2	0.15	605	0.7	0.45	0.55	0.15	0.05	0.25
206	0.05	0.7	0	0.8	0	0	408	0.7	0.85	0.7	0.25	0	0
209	0.05	0.8	0.05	0.95	0	0	113	0.7	0.95	0.6	0.1	0	0
305	0.05	0.85	0.05	0.95	0.05	0.05	207	0.75	0.65	0.6	0	0.05	0.05
409	0.05	0.9	0	0.95	0.25	0	309	0.75	0.9	0.55	0	0.75	0.05
413	0.05	1	0.05	0.95	1	0	410	0.8	0.05	0.65	0.1	0	0.95
505	0.1	0	0	0	0	0.9	303	0.8	0.25	0.8	0.2	0.25	0.75
402	0.1	0.05	0	0.75	0.05	0.95	215	0.8	0.55	0.75	0.2	0.45	0.3
623	0.1	0.25	0.05	0.8	0.25	0.75	405	0.8	0.75	0.2	0.1	0	0
415	0.1	0.85	0.05	0.85	0.4	0	602	0.85	0	0.85	0.15	0	0.85
304	0.15	0	0.05	0.75	0	0.95	302	0.85	0.3	0.85	0.15	0.3	0.7
219	0.15	0	0.15	0.85	0	0.85	220	0.9	0.2	0.9	0.1	0.2	0.8
105	0.15	0.8	0.15	0.85	0.55	0.15	107	0.95	0.25	0.95	0.05	0.05	0.55
518	0.15	0.95	0.05	0.7	0.65	0	618	1	0.35	0.7	0	0.35	0.6
315	0.15	1	0.15	0.85	0.15	0	106	1	0.5	1	0	0.4	0.5
308	0.2	0.05	0.2	0.8	0.05	0.95	112	1	0.55	1	0	0.35	0.35
110	0.2	0.4	0.2	0.65	0	0.3	401	1	0.8	1	0	0.8	0.2
204	0.2	0.4	0.2	0.8	0.1	0.5	510	1	1	0	0	1	0
101	0.2	0.7	0	0.45	0.25	0.1	506	1	1	1	0	1	0
514	0.2	0.8	0.2	0.8	0.3	0							

doi:10.6342/NTU202002331

4 Conclusion

In this experiment we set out to examine how people process potentially irrelevant information when they already established certain pre-existing beliefs. To uncover the mechanism behind confirmation bias, we ask subjects to report beliefs of the assigned urn, in which they have prior beliefs and a piece of potentially irrelevant information. Crucially, they also have to report beliefs of the irrelevant urn, by which we can visually observe the strength of weight they put on the potentially irrelevant information. We show that subjects tend to view this information as completely worthless in evaluating the assigned urn when it conflicts their prior beliefs, but overvalue it when it confirms their prior beliefs. We estimate the tendency of attributing the information to the irrelevant urn. The results suggest that on average subjects believe the information is from the irrelevant urn with probabilities more regardless of the types of information. However, they increase the probabilities when the information is conflicting by 12%. When we allow subjects consider other's information might be inaccurate, the they still believe the information is more likely from the irrelevant urn when it is conflicting. These results are robust even we assume subjects independently make decisions on the assigned and irrelevant urn.

Most importantly, we try to explore the mechanism leading to the echo chamber, especially focusing on the information updating. By explicitly creating an irrelevant urn, we highlight one possible reason people usually stick to their political stance or beliefs on controversial issues, even leading to polarization. Though this may not be the only cause of the echo chamber effect, our results suggest that dismissing the information when it conflicts with one's prior is still a prominent cause.

doi:10.6342/NTU202002331

References

L. Babcock, G. Loewenstein, S. Issacharoff, and C. Camerer. Biased judgments of fairness in bargaining. *American Economic Review*, 85(5):1337–1343, 1995.

M. A. Costa-Gomes and V. P. Crawford. Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review*, 96(5):1737–1768, 2006.

A. Coutts. Good news and bad news are still news: Experimental evidence on belief updating. *Experimental Economics*, 22(2):369–395, 2019.

D. Eil and J. M. Rao. The good news-bad news effect: asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–38, 2011.

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

R. G. Fryer Jr, P. Harms, and M. O. Jackson. Updating beliefs when evidence is open to interpretation: Implications for bias and polarization. *Journal of the European Economic Association*, 17(5):1470–1501, 2019.

E. L. Glaeser and C. R. Sunstein. Why does balanced news produce unbalanced views? Technical report, National Bureau of Economic Research, 2013.

B. Greiner. Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association*, 1(1):114–125, 2015.

doi:10.6342/NTU202002331

D. M. Grether. Bayes rule as a descriptive model: The representativeness heuristic. *The Quarterly Journal of Economics*, 95(3):537–557, 1980.

C. A. Holt and A. M. Smith. An update on bayesian updating. *Journal of Economic Behavior & Organization*, 69(2):125–134, 2009.

C. A. Holt and A. M. Smith. Belief elicitation with a synchronized lottery choice menu that is invariant to risk attitudes. *American Economic Journal: Microeconomics*, 8(1):110–39, 2016.

D. M. Kahan, H. Jenkins-Smith, and D. Braman. Cultural cognition of scientific consensus. *Journal of Risk Research*, 14(2):147–174, 2011.

D. M. Kahan, E. Peters, M. Wittlin, P. Slovic, L. L. Ouellette, D. Braman, and G. Mandel. The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, 2(10):732–735, 2012.

C. G. Lord, L. Ross, and M. R. Lepper. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37(11):2098, 1979.

A. Tversky and D. Kahneman. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2):207–232, 1973.

doi:10.6342/NTU202002331

Appendix

A First Phase Belief

The data points aligned with 45 degree line in the irrelevant urn, implying that subjects believe the initial draw can infer both urns. Figure 9a shows that a majority of these choices are made by different subjects and they only perform this behavior one time. Moreover, Figure 9b shows the occurred round of these choices. They do not concentrate on particular rounds, suggesting that such unusual behavior is randomly made throughout the experiment and is unlikely explained by learning effect.

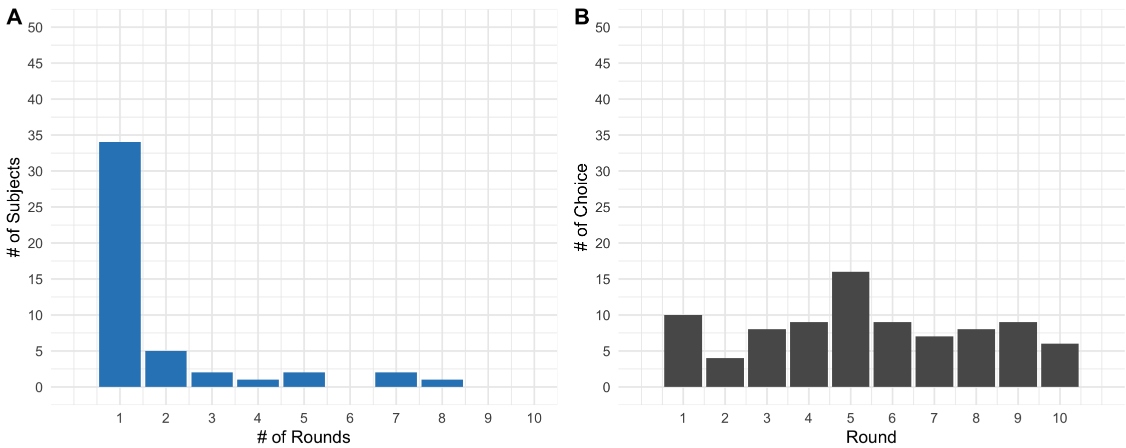


Figure 9: Beliefs Aligned with 45 Degree Line in the Irrelevant Urn. (a) the Number of Rounds (b) Occurrence Rounds.

B Second Phase Raw Data

Figure 10 shows the raw data of second phase beliefs. In particular, it is clear to see the overreaction in the irrelevant urn.

doi:10.6342/NTU202002331

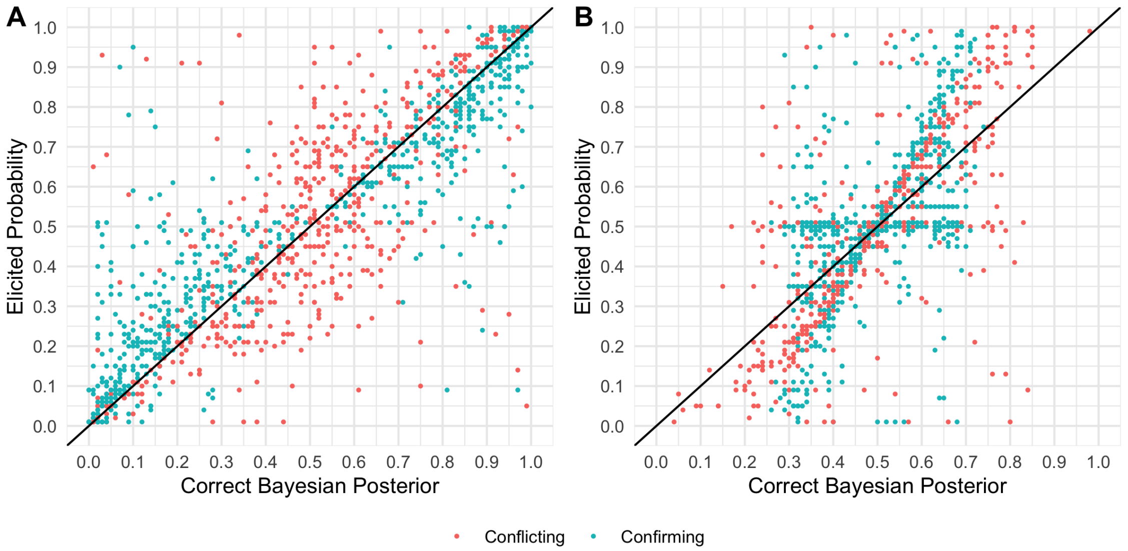


Figure 10: Elicited Beliefs in the Second Phase of the (a) Assigned (b) Irrelevant Urn

C Alternative Experimental Designs

We document alternative designs that were eventually dropped. Our first experimental design is inspired by Eil and Rao (2011). Subjects are asked to predict the real value of an asset with ten possible states. The computer randomly draws with replacement three balls from twelve, in which ten balls represent the ten possible states and the additional two balls represent the real value. Thus, the real value is drawn with probability 0.25 compared to others with 0.083. After observing their private information of three ball draws, they report their beliefs of each state that add up to 1.

Subjects then observe new information: The computer divides others into two halves, one half whose predictions are close to and the other half whose predictions are far from the subject, and randomly draws another subject from one of them to reveal his/her prediction. The procedure is repeated three times, so three other subjects' predictions will be revealed to the subject. We elicit beliefs in terms of

doi:10.6342/NTU202002331

probabilities after subjects observe each piece of information using the quadratic scoring rule. The experimental interface is shown in Figure 11.

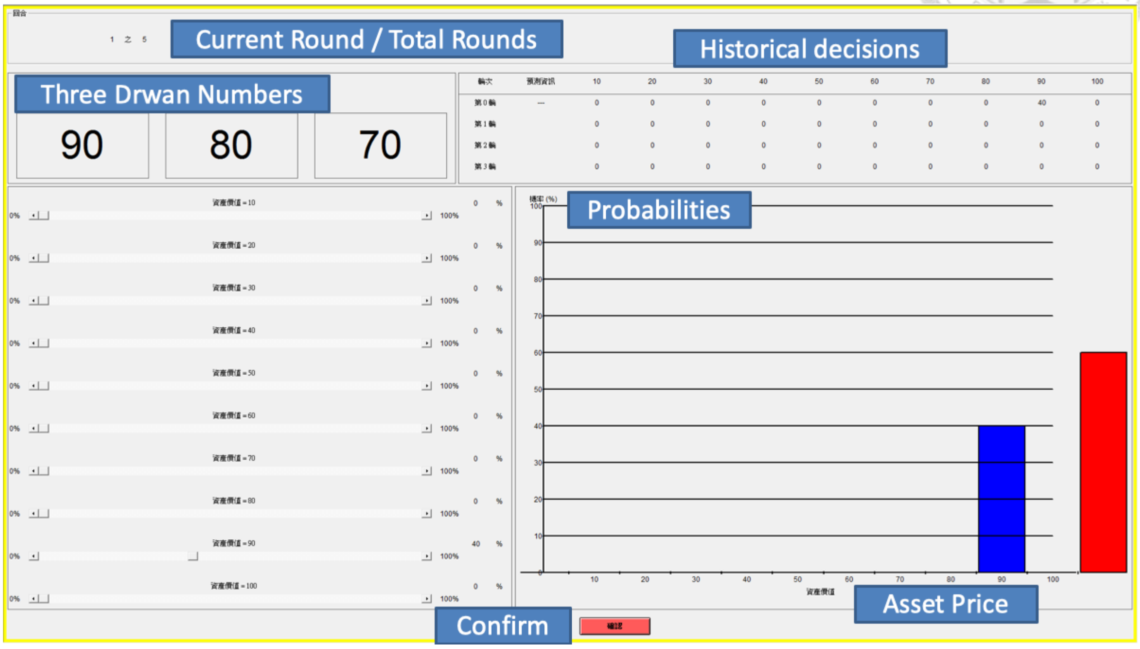


Figure 11: Screen Shot of the First Version Experiment.

Our second experimental design is similar to the first one, but with only two possible states. There are two urns, A and B, in the experiment. Urn A applies the *Maximum Rule* and Urn B applies the *Minimum Rule*, so each urn reports either maximum or minimum of two draws from the uniform distribution. We provide the probability table in case subjects cannot figure it out themselves. Subjects observe a ball from urn A or B with equal chance, and report the probability that the chosen urn is A. Then, subjects observe others' information and beliefs are elicited using the same design as the first version.

Our third experimental design is nearly identical to our final one implemented, but with three important differences. First of all, it is a one shot game with three stages of belief-updating, while the final experiment has ten rounds each with one

doi:10.6342/NTU202002331

stage of belief-updating. In other words, subjects observe their initial draw and then receive three other piece of information. Second, we use the BDM procedure as in Coutts (2019) to elicit beliefs, which is illustrated in Figure 12a. Finally, the probability for drawing each number under the *Maximum Rule* and *Minimum Rule* is shown in tables. The experimental interface is shown in Figure 12b.

doi:10.6342/NTU202002331

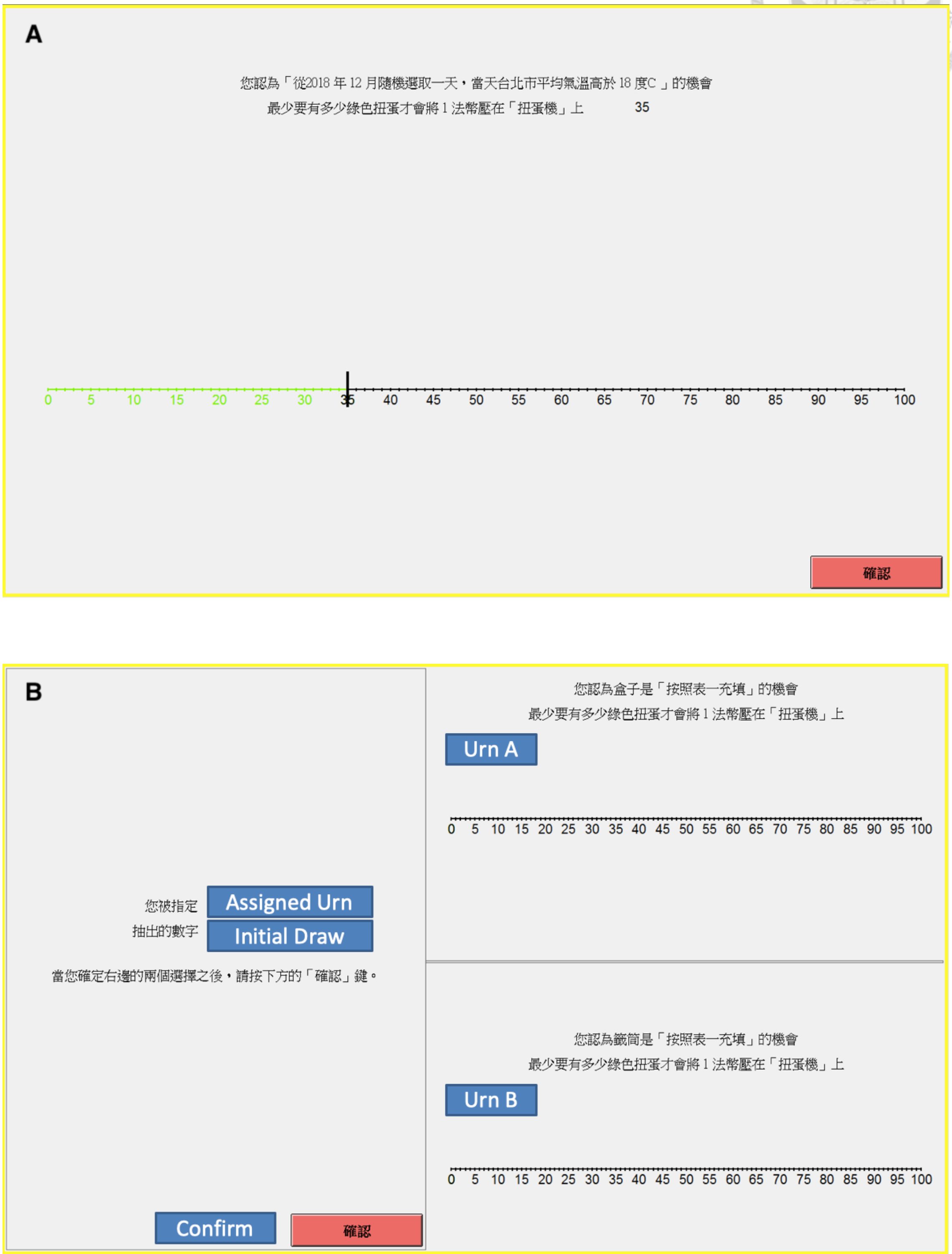


Figure 12: Screen Shot of Third Version Experiment.

doi:10.6342/NTU202002331

TASSEL 實驗說明 p.1

實驗報酬

感謝您參加本實驗。本實驗結束後，您將得到定額車馬費新台幣 200 元，以及您在實驗中獲得的「法幣」所兌換成之新台幣。（「法幣」為本實驗的實驗貨幣單位。）您在實驗中能獲得的「法幣」會根據您所做的決策、別人的決策，以及隨機亂數決定，每個人都不同。每個人都會個別獨自領取報酬，您沒有義務告訴其他人您的報酬多寡。

請注意：本實驗法幣與新台幣的兌換匯率為 1 : 50，即 1 法幣＝50 元新台幣。

實驗說明—第一部分

在今天的實驗中，您需要回答若干個「您認為某事件發生的機會」。為了讓您瞭解如何回答這些問題，我們在第一部分中，提供兩個問題給大家練習，並且，這兩個練習都會有報酬。我們選取世界上兩個城市，請您思考以下問題：

1. 從 2018 年 12 月隨機選取一天，當天台北市平均氣溫高於 18 度 C 的機會。
2. 從 2018 年 12 月隨機選取一天，當天杜尚貝市平均氣溫高於 18 度 C 的機會。

我們以第一個問題為例，以下詳細說明在實驗中如何回答上述問題。

<u>情境壓寶</u>

在每個問題中，實驗者會給每位受試者 1 法幣做選擇。您必須決定這 1 法幣要壓在以下哪一種情境：

1. 「台北市」：在 2018 年 12 月隨機選取的那一天，台北市平均氣溫高於 18 度 C。
2. 「扭蛋機」：從扭蛋機中隨機抽出一個扭蛋是「綠色」。

如果情境為真，則您將得到這 1 法幣的報酬；若否，則沒有報酬。也就是說：

1. 若您決定將這 1 法幣壓在「台北市」，若在 2018 年 12 月隨機選取的那一天，台北市平均氣溫高於 18 度 C，您將得到這 1 法幣的報酬；若否，則沒有報酬。
2. 若您決定將這 1 法幣壓在「扭蛋機」，電腦會隨機從扭蛋機中抽出一個扭蛋，若為「綠色」，您將得到這 1 法幣的報酬；若否，則沒有報酬。

doi:10.6342/NTU202002331

TASSEL 實驗說明 p.2

扭蛋機中共有 100 個扭蛋，其中若干個是綠色扭蛋，剩下是黑色扭蛋。因此，如果扭蛋機有 40 個綠色扭蛋和 60 個黑色扭蛋，若從其中抽出一個，抽到綠色扭蛋的機會為 40%。因此，您需要仔細思考您認為從 2018 年 12 月隨機選取一天，當天台北市平均氣溫高於 18 度 C 的機會是否高於或低於 40%。

如果您覺得從 2018 年 12 月隨機選取一天，當天台北市平均氣溫高於 18 度 C 的機會為三分之一：

1. 當綠色扭蛋數小於或等於 33，壓在「台北市」有較大機會得到 1 法幣。
2. 當綠色扭蛋數大於 34，壓在「扭蛋機」有較大機會得到 1 法幣。

如果您覺得從 2018 年 12 月隨機選取一天，當天台北市平均氣溫高於 18 度 C 的機會為 70%：

1. 當綠色扭蛋數小於或等於 69，壓在「台北市」有較大機會得到 1 法幣。
2. 當綠色扭蛋數大於 71，壓在「扭蛋機」有較大機會得到 1 法幣。
3. 當綠色扭蛋數等於 70，壓在「扭蛋機」或「台北市」機會相同。

<u>要多少綠色扭蛋才要壓在「扭蛋機」</u>

<u>在尚未被告知扭蛋機有多少個綠色扭蛋之前，您需要先決定最少有多少個綠色扭蛋，您才會願意將這 1 法幣壓在「扭蛋機」上。</u>然後，電腦會隨機放入 0 到 100 個綠色扭蛋進入扭蛋機，每種可能的機會相同，與您的決定無關。也就是說，扭蛋機中有 0 個綠色扭蛋、1 個綠色扭蛋、2 個綠色扭蛋、……、100 個綠色扭蛋的機會都相同。

最後，根據您決定的最少數量，實驗者會幫您將這 1 法幣壓在機會較高的那個情境（如果機會相同，則壓在「扭蛋機」）：若扭蛋機中的綠色扭蛋少於您願意壓在「扭蛋機」的最少數量，則這 1 法幣會壓在「台北市」上；若扭蛋機中的綠色扭蛋多於（或等於）您願意壓在「扭蛋機」的最少數量，則這 1 法幣會壓在「扭蛋機」上。

doi:10.6342/NTU202002331

TASSEL 實驗說明 p.3

以下舉一個例子詳細說明：若您決定扭蛋機中最少需要有 35 個綠色扭蛋，您才願意將這 1 法幣壓在「扭蛋機」上，則可能的情況有以下兩種：

1. 電腦放入至少 35 個綠色扭蛋（35、36、……、100）：此時實驗者會幫您將這 1 法幣壓在「扭蛋機」上。如果電腦隨機從扭蛋機抽出的是綠色扭蛋，則您將獲得 1 法幣的報酬；若否，則不會獲得報酬。

2. 電腦放入少於 35 個綠色扭蛋（34、33、……、0）：此時實驗者幫您將這 1 法幣壓在「台北市」上，若在 2018 年 12 月隨機選取的那一天，台北市平均氣溫高於 18 度 C，則您將獲得 1 法幣的報酬；若否，則不會獲得報酬。

請注意，在這個設計下，您選擇綠色扭蛋的最少數量，在以下情況對您是最有利的：誠實且精確地按照您覺得從 2018 年 12 月隨機選取一天，當天台北市平均氣溫高於 18 度 C 的機會高低來回答。

以上用如何回答「從 2018 年 12 月隨機選取一天，當天台北市平均氣溫高於 18 度 C 的機會」來說明，回答「從 2018 年 12 月隨機選取一天，當天杜尚貝市平均氣溫高於 18 度 C 的機會」的方式是一樣的。

第一部分的實驗是為了讓您瞭解如何在接下來的正式實驗中回答其他問題，因此，如果您對這個部分的實驗有任何疑問，請現在舉手，實驗者會過來為您解答。並且，在實驗開始後，若您在回答這兩個問題的過程中有任何不清楚的地方，請當場舉手，實驗者會過來為您解答。

doi:10.6342/NTU202002331

TASSEL 實驗說明 p.4

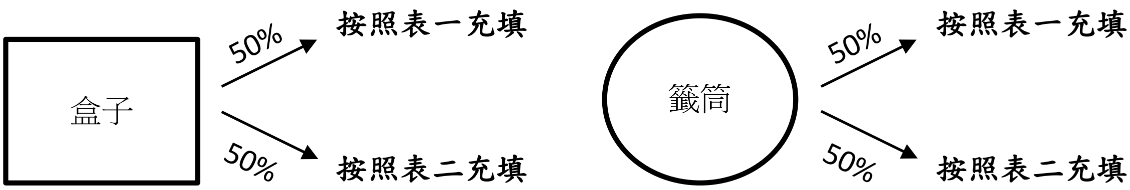
實驗說明—第二部分

以下將開始正式實驗，本實驗中有一個盒子、一個籤筒。盒子要充填 10,000 顆球，籤筒要充填 10,000 支籤。每顆球/每支籤上都標示著一個「一」到「一百」的數字。盒子和籤筒充填時，球/籤的組成有以下兩種可能：

按照表一充填：寫著數字「一」的球/籤有 1 顆、寫著數字「二」的球/籤有 3 顆、寫著數字「三」的球/籤有 5 顆、……、寫著數字「九十九」的球/籤有 197 顆、寫著數字「一百」的球/籤有 199 顆，總共 10,000 顆。

按照表二充填：寫著數字「一」的球/籤有 199 顆、寫著數字「二」的球/籤有 197 顆、寫著數字「三」的球/籤有 195 顆、……、寫著數字「九十九」的球/籤有 3 顆、寫著數字「一百」的球/籤有 1 顆，總共 10,000 顆。

盒子和籤筒是按照表一或表二充填的機會都相等，且盒子和籤筒是各自分別充填，互不影響。也就是說，電腦會先擲一枚公正的十元銅板決定盒子如何充填，若為正面(50%)，則盒子會按照表一來充填；若為反面(50%)，則盒子會按照表二來充填。然後電腦會另外擲一枚公正的一元銅板決定籤筒如何充填，若為正面(50%)，則籤筒會按照表一來充填；若為反面(50%)，則籤筒會按照表二來充填。在第二部分的一開始，電腦就會先決定盒子和籤筒的充填方式，此充填方式在整場實驗中不會改變，並且，所有參與同學都會面對同樣的一個盒子和同樣的一個籤筒。



在電腦決定完盒子和籤筒是按照表一或表二充填後，電腦會指定您從盒子抽球或從籤筒抽籤，您被指定從盒子中抽球的機會為 50%；同樣地，您被指定從籤筒中抽籤的機會也是 50%。電腦在指定所有參與同學從盒子抽球或從籤筒抽籤的過程是獨立的，也就是說，您被指定從盒子抽球或從籤筒抽籤的結果，並不會影響到其他同學被指定從盒子抽球或從籤筒抽籤的機會。

如果您被指定從盒子中抽球，您不會獲得任何跟籤筒有關的資訊；如果您被指定從籤筒中抽籤，您不會獲得任何跟盒子有關的資訊。電腦會先告知您被指定到哪一個(盒子或籤筒)，再隨機從盒子或籤筒中抽出一個給您看。

doi:10.6342/NTU202002331

TASSEL 實驗說明 p.5

當您看到所抽出的數字後，您要做出兩個選擇，一個跟盒子有關，另一個跟籤筒有關：

1. 盒子是「按照表一充填」的機會。
2. 籤筒是「按照表一充填」的機會。

這兩個選擇各自獨立，因為電腦是先擲一枚公正的十元銅板決定盒子如何充填，然後再另外擲一枚公正的一元銅板決定籤筒如何充填。

回答以上兩個選擇的方式和第一部分完全相同，我們用回答跟盒子有關的選擇再次說明重點，回答跟籤筒有關的選擇是一樣的。在每個選擇中，實驗者會給每位受試者 1 法幣。這 1 法幣會壓在以下其中一種情境：

1. 「盒子」：盒子是「按照表一充填」的。
2. 「扭蛋機」：從扭蛋機中隨機抽出一個扭蛋是「綠色」。

如果情境為真，則您將得到這 1 法幣的報酬；若否，則沒有報酬。

扭蛋機中共有 100 個扭蛋，其中若干個是綠色扭蛋，剩下是黑色扭蛋。**在尚未被告知扭蛋機有多少個綠色扭蛋之前，您需要先決定最少有多少個綠色扭蛋，您才會願意將這 1 法幣壓在「扭蛋機」上。**然後，電腦會隨機放入 0 到 100 個綠色扭蛋進入扭蛋機，每種可能的機會相同，與您的決定無關。也就是說，扭蛋機中有 0 個綠色扭蛋、1 個綠色扭蛋、2 個綠色扭蛋、……、100 個綠色扭蛋的機會都相同。

最後，根據您決定的最少數量，實驗者會幫您將這 1 法幣壓在機會較高的那個情境（如果機會相同，則壓在「扭蛋機」)：若扭蛋機中的綠色扭蛋少於您願意壓在「扭蛋機」的最少數量，則這 1 法幣會壓在「盒子」上；若扭蛋機中的綠色扭蛋多於（或等於）您願意壓在「扭蛋機」的最少數量，則這 1 法幣會壓在「扭蛋機」上。

請注意，在這個設計下，您選擇綠色扭蛋的最少數量，在以下情況對您是最有利的：誠實且精確地按照您覺得盒子是「按照表一充填」的機會高低來回答。

以上用「跟盒子有關的選擇」來說明，「跟籤筒有關的選擇」是一樣的，但是這兩個選擇各自獨立，因為電腦是先擲一枚公正的十元銅板決定盒子如何充填，然後再另外擲一枚公正的一元銅板決定籤筒如何充填。

如果您對這個部分的實驗有任何疑問，請現在舉手，實驗者會過來為您解答。

doi:10.6342/NTU202002331

TASSEL 實驗說明 p.6

實驗說明—第三部分

第三部分的實驗中，電腦會隨機抽出其他三位不同的參與同學，並依序將這三位同學在第二部分的其中一個選擇給您看。

請注意，在第二部分的實驗中，每位同學都做了兩個選擇，一個跟盒子有關，另一個跟籤筒有關。電腦只會給您看該位同學的其中一個選擇，並且不會告知您這個選擇是跟盒子有關還是跟籤筒有關。如果電腦指定該位同學從盒子中抽球，我們會給您看他跟盒子有關的選擇。如果電腦指定該位同學從籤筒中抽籤，則會給您看他跟籤筒有關的選擇。

當您看完每一位隨機抽出的參與同學在第二部分的其中一個選擇後，您都需要重新再做一次兩個選擇，一個跟盒子有關，另一個跟籤筒有關：

1. 盒子是「按照表一充填」的機會。
2. 籤筒是「按照表一充填」的機會。

以上兩個選擇的方式和第一、二部分完全相同。

因此，在第三部分的實驗中，實驗者總共會給您 6 法幣，當您看完第一位隨機抽出的參與同學在第二部分的其中一個選擇後，您將需要用 1 法幣做出一個跟盒子有關的選擇；同樣地，您也需要用 1 法幣做出一個跟籤筒有關的選擇。再做完這兩個選擇後，您才會看到第二位隨機抽出的參與同學在第二部分的其中一個選擇，並再次作出一個跟盒子有關的選擇及跟籤筒有關的選擇。依此類推。

如果您對這個部分的實驗有任何疑問，請現在舉手，實驗者會過來為您解答。

doi:10.6342/NTU202002331

表一：每種數字有幾顆球/幾支籤（總共 10,000 顆球/支籤）

數字	顆/支	數字	顆/支	數字	顆/支	數字	顆/支
一	1	二十六	51	五十一	101	七十六	151
二	3	二十七	53	五十二	103	七十七	153
三	5	二十八	55	五十三	105	七十八	155
四	7	二十九	57	五十四	107	七十九	157
五	9	三十	59	五十五	109	八十	159
六	11	三十一	61	五十六	111	八十一	161
七	13	三十二	63	五十七	113	八十二	163
八	15	三十三	65	五十八	115	八十三	165
九	17	三十四	67	五十九	117	八十四	167
十	19	三十五	69	六十	119	八十五	169
十一	21	三十六	71	六十一	121	八十六	171
十二	23	三十七	73	六十二	123	八十七	173
十三	25	三十八	75	六十三	125	八十八	175
十四	27	三十九	77	六十四	127	八十九	177
十五	29	四十	79	六十五	129	九十	179
十六	31	四十一	81	六十六	131	九十一	181
十七	33	四十二	83	六十七	133	九十二	183
十八	35	四十三	85	六十八	135	九十三	185
十九	37	四十四	87	六十九	137	九十四	187
二十	39	四十五	89	七十十	139	九十五	189
二十一	41	四十六	91	七十一	141	九十六	191
二十二	43	四七	93	七二	143	九十七	193
二十三	45	四八	95	七三	145	九十八	195
二十四	47	四九	97	七四	147	九十九	197
二十五	49	五十	99	七五	149	一百	199

doi:10.6342/NTU202002331

表二：每種數字有幾顆球/幾支籤（總共 **10,000** 顆球/支籤）

數字	顆/支	數字	顆/支	數字	顆/支	數字	顆/支
一	199	二十六	149	五十一	99	七十六	49
二	197	二十七	147	五十二	97	七十七	47
三	195	二十八	145	五十三	95	七十八	45
四	193	二十九	143	五十四	93	七十九	43
五	191	三十	141	五十五	91	八十	41
六	189	三十一	139	五十六	89	八十一	39
七	187	三十二	137	五十七	87	八十二	37
八	185	三十三	135	五十八	85	八十三	35
九	183	三十四	133	五十九	83	八十四	33
十	181	三十五	131	六十	81	八十五	31
十一	179	三十六	129	六十一	79	八十六	29
十二	177	三十七	127	六十二	77	八十七	27
十三	175	三十八	125	六十三	75	八十八	25
十四	173	三十九	123	六十四	73	八十九	23
十五	171	四十	121	六十五	71	九十	21
十六	169	四十一	119	六十六	69	九十一	19
十七	167	四十二	117	六十七	67	九十二	17
十八	165	四十三	115	六十八	65	九十三	15
十九	163	四十四	113	六十九	63	九十四	13
二十	161	四十五	111	七十	61	九十五	11
二十一	159	四十六	109	七十一	59	九十六	9
二十二	157	四十七	107	七十二	57	九十七	7
二十三	155	四十八	105	七十三	55	九十八	5
二十四	153	四十九	103	七十四	53	九十九	3
二十五	151	五十	101	七十五	51	一百	1

doi:10.6342/NTU202002331