ソシオネットワーク戦略ディスカッションペーパーシリーズ ISSN 1884-9946
第 62 号　2018 年 6 月
RISS Discussion Paper Series
No.62　June, 2018

Supplementary Results for "Meaningful Learning in Weighted Voting Games: An Experiment"

Naoki Watanabe

文部科学大臣認定　共同利用・共同研究拠点

関西大学ソシオネットワーク戦略研究機構

Research Institute for Socionetwork Strategies,
Kansai University
Joint Usage / Research Center, MEXT, Japan
Suita, Osaka, 564-8680, Japan
URL: https://www.kansai-u.ac.jp/riss/index.html
e-mail: riss@ml.kandai.jp
tel. 06-6368-1228
fax. 06-6330-3304

Supplementary Results for "Meaningful Learning in Weighted Voting Games: An Experiment"*

Naoki Watanabe†

November 16, 2021

Abstract

This paper clarifies subjects' search behavior of correct options behind the experimental results shown by Guerci et al. (2017). In the experiment, subjects were asked to choose one of two weighted voting games repeatedly and their payoffs are stochastically determined for each of their choice according to a payoff-generating function that was hidden from subjects. The main results are as follows. (1) In the additional sessions conducted for the treatment without any payoff-related feedback information, it was reconfirmed that subjects learned to choose the correct option that generates higher expected payoffs for them and generalized what they had thought introspectively in a binary choice problem to a similar but different one. (2) Feedback information about payoffs given immediately after subjects' choice often confused their inference on the relationship between nominal voting weights and actual payoffs so that they took the win-stay-lose-shift strategy in some sessions. (3) Immediate payoff-related feedback information sometimes induced subjects to randomly choose the runs of options.

Keywords: meaningful learning, bandit experiment, weighted voting, search behavior, win-stay-lose-shift strategy

JEL Classification: C91, D72, D83

*This paper is a revised version of Kansai University RISS Discussion Paper Series No. 62 issued in 2018. The author wishes to thank Eric Guerci and Nobuyuki Hanaki, for giving him their permission to reuse the data, and Daniel Friedman, Tetsuya Kawamura, Midori Hirokawa, Yoichi Hizen, Tatsuya Kameda, Naoko Nishimura, and Kazuhito Ogawa for their comments on that previous work which motivated him to provide these supplementary results. This research could not have been completed without the research assistance of Yoichi Izunaga, Toru Suzuki, Hiroshi Tanaka, and Yosuke Watanabe. Financial support from Foundation for the Fusion of Science and Technology (FOST), MEXT Grants-in-Aid 24330078 and 25380222, JSPS-ANR bilateral research grant "BECOA" (ANR-11-FRJA-0002), and Joint Usage/Research of ISER at Osaka University are gratefully acknowledged.

†Graduate School of Business Administration, Keio University, 4-1-1 Hiyoshi Kohoku, Yokohama, Kanagawa 223-8526, Japan; E-mail: naoki50@keio.jp; Phone: +81-45-564-2039

1 Introduction

It would be truly difficult for people to deal with weighted voting; Felsenthal and Machover (1998, pp.164-165) noted that it might be difficult even for the policy makers and officials who designed and re-designed the system to see the latent relationship between the actual voting powers and the nominal voting weights, and Gelman et al. (2004) stated that the standard theoretical indices of voting power do not predict the actual voting outcomes. Weighted voting is, however, a popular collective decision-making system. In order for us to use this system better, it is important to examine whether people can learn from their experiences the underlying structure of weighted voting.

When people generalize what they have learned in a situation to a similar but different one, that higher order concept of learning is called meaningful learning (Rick and Weber (2010)).[1] In a two-armed bandit experiment, Guerci et al. (2017) could not observe meaningful learning of subjects when immediate payoff-related feedback information was given to them, whereas they observed it only in sessions for the treatment without any payoff-related feedback information. Feedback information is essentially incorporated also in the standard theory of reinforcement learning (e.g., Erev and Roth, 1998), belief-based learning (e.g., Cheung and Friedman, 1997), and experience weighted attraction learning (Camerer and Ho, 1999). What hindered the human subjects from meaningful learning when immediate feedback information was given to them? For taking a new step to the future research, this paper clarifies some features on subjects' search behavior behind the experimental results shown by Guerci et al. (2017), reexamining the data in their sessions and additional sessions.

In strategic situations, feedback information provided to each subject contains the outcomes generated by unplanned or exploratory behavior of other subjects, and thus individual inferences might be confused mutually among subjects. Guerci et al. (2017) thus drastically simplified the experimental design to remove subjects' learning through their strategic interaction. The experimental design is as follows. In each session subjects choose one of two weighted voting games (options) repeatedly and obtain their payoffs which are stochastically generated for each choice they make according to a voting theory. The binary choice problems are different between the first and second parts of the session, but the payoff generating function in binary choice problems remains the same. Subjects thus have a chance to learn something underlying the situation they face in the first part and to apply what they learned in the first part to their decision in the second part.

[1] Meaningful learning is also called "transfer of learning" (Cooper and Kagel, 2003, 2008) or "epiphany" (Dufwenberg et al., 2010).

There are three treatments; no feedback, partial feedback, and full feedback. For the no-feedback treatment, subjects are not informed of any payoffs they receive as the result of their choice until the session ends. For the partial-feedback treatment, each subject is informed of his or her own payoff. For the full-feedback treatment, after each choice subjects are informed of the payoffs including others. Intuitively, for each binary choice problem, it is said that meaningful learning is observed in the problem when there is a significant increase in number of subjects who chose the correct option between subjects who experienced similar but different problem and subjects who did not.

Guerci et al. (2017) observed that (i) subjects learned to choose the correct option that generates higher expected payoffs for them even without any payoff-related feedback information and that (ii) in each sequence that starts with an easy binary choice problem and then gives a difficult one there was statistically significant evidence of meaningful learning only in sessions for the treatment with no payoff-related feedback information. In their sessions, however, meaningful learning with no feedback information might be a consequence of inertia on subjects' choice, because the correct options were the same choice (Choice 2) in the first and second parts of the sequences of binary choice problems. We should thus confirm whether observation (ii) was not the consequence of the inertia on subjects' choice by changing the correct options between the two parts.

Hypothesis 1. *Meaningful learning without any payoff-related feedback information is not the consequence of the inertia on subjects' choice, and it is observed in each sequence that starts with an easy binary choice problem and then gives a difficult one.*

Another major question is, as stated above, is to investigate what hindered subjects from gaining a deeper insight on the underlying structure of weighted voting when immediate feedback information was given to them. In the post-experimental questionnaire, some subjects reported that they changed their choices when they received zero points, while they did not when they received positive points. It would be plausible that subjects took this "win-stay-lose-shift" strategy (Nowak and Sigmund, 1993).

Hypothesis 2. *Subjects took the win-stay-lose-shift strategy, when feedback information is provided immediately after their choice.*

Together with Hypothesis 2, we should confirm whether they searched for the correct options with some certain rule, when they were given immediate feedback information.

Hypothesis 3. *Immediate payoff-related feedback information induced subjects to randomly choose the runs of options.*

This paper shows the following results, reexamining the data used in Guerci et al. (2017) and additional sessions conducted by the author. (1) It was reconfirmed that without any feedback information, subjects meaningfully learned to choose the correct option in one of two sequences in which Guerci et al. (2017) observed subjects' meaningful learning, even though the correct option was changed from Choice 1 in the first part to Choice 2 in the second part in the sequence. The answers to the post-experimental questionnaire show their proper reasoning of their choice. Accordingly, meaningful learning was not observed by chance. (2) For treatments with immediate feedback information, it was confirmed that subjects took the win-stay-lose-shift strategy in some sessions, but (3) in other sessions they did not search for the correct options with any certain rules.

In particular, we found the following search behavior of the subjects in a binary choice problem where meaningful learning was observed in the treatment with no payoff-related feedback information. Subjects chose their options in almost the same thinking time, regardless of the treatments in which they participated. In the partial-feedback treatment, subjects did not search for the correct options with any certain rules. In the full-feedback treatment, experienced subject took the win-stay-lose-shift strategy. There was no significant difference in their thinking time among all treatments. Accordingly, we do not have to care about the effect of thinking time on their behavior in the comparison.

In summary, Hypothesis 1 was affirmatively confirmed and Hypotheses 2 and 3 were partly confirmed. Result (1) was not obtained by chance. In the post experimental questionnaire, subjects noted the proper reasoning for their choice. According to Results (2) and (3), subjects' inference on the relationship between nominal voting weights and actual payoffs might be mistaken or confused when payoff-related feedback information was given immediately after their choice. Withholding such misleading or confusing information allows subjects to deliberate the underlying structure of weighted voting.

The rest of this paper proceeds as follows. Section 2 describes the experimental design. Sections 3 provides observations and Result (1) on subjects' learning and meaningful learning. Section 4 presents Results (2) and (3), where we deal with the length of time subjects spent for thinking, win-stay-lose-shift strategy, and their random choice of runs in each sequence of the options. Section 5 notes two brief remarks for further research, based on our results and observations. One is the verification of implications derived by theoretical models, and the other is the relationship between subjects' cognitive ability and their meaningful learning. Appendix A provides the instructions of the experiment, and Appendix B summarizes the comments of subjects who participated in the sessions for Guerci et al. (2017) for reference.

2 Experimental Design

This section describes the experimental design. Subsection 2.1 explains the session outline, Subsection 2.2 shows what subjects see on their monitors in the session, and Subsection 2.3 refers to the experimental procedure. The instructions given to subjects are attached in Appendix A, although the instructions for bandit experiments are extremely simple.

2.1 Session Outline

There are three treatments, which are explained later at the end of this subsection. Each session in those treatments consists of 60 periods. At each period, the subject is asked to choose one of two committees of four members who are supposed to divide 120 points among them. Let $N = \{1, 2, 3, 4\}$ be the set of the members (players). A committee (a weighted voting game) is represented by $[q; v_1, v_2, v_3, v_4]$, where q is the quota (the minimum number of votes required for an allocation to be adopted) and v_i is the voting weight (the number of votes) allocated to player $i \in N$. Both committees have the same numbers of total votes and quota as well as the same number of votes for player 1. Every subject acts as player 1 and faces a binary choice problem for the first 40 periods, while in the following 20 periods he or she faces another but different one; for example, in the first 40 periods subjects face a choice between $[14; 5, 3, 7, 7]$ and $[14; 5, 4, 6, 7]$, while in the following 20 periods they are faced with $[6; 1, 2, 3, 4]$ and $[6; 1, 1, 4, 4]$. Subjects are not asked to play the weighted voting games that they choose.[2] In the instruction, subjects are clearly informed that the other members of the committees are all fictitious. This experiment is thus regarded as a two-armed bandit experiment with contextual information on weighted voting.

The payoff each subject obtains from his or her choice is externally determined according to a power index called DPI (Deegan and Packel, 1978), which is defined as follows. Given a weighted voting game, a non-empty subset S of N is called a coalition, and a coalition is called a winning coalition if $\sum_{i \in S} v_i \geq q$; otherwise, it is called a losing coalition. A minimum winning coalition (MWC) is a winning coalition such that deviation by any member of the coalition turns its status from winning to losing. In the experiment, for each period, one MWC is drawn with equal probability from all the possible MWCs for the committee that the subject has chosen. If the subject is a member of the drawn MWC, then he or she receives an equal share of the total payoff with the other members.

[2] This setting was made to avoid the complexities mentioned at the beginning of the Introduction; subjects are simultaneously learning to play a weighted voting game from interfering with the other subjects learning about the underlying relationship between their nominal voting weights and their expected payoffs.

The binary choice problems we use are listed in Table 1. We denote each MWC by the votes apportioned to its members; the MWCs in a committee $[14; 5, 3, 7, 7]$ are written as $(5, 3, 7)$, $(5, 3, 7)$, and $(7, 7)$. When $[14; 5, 3, 7, 7]$ is chosen by a subject (player 1), player 1 has a 2/3 chance of being on the MWCs and will receive 1/3 of the total payoff (120 points) for being on each of the MWCs, while there is a 1/3 probability that player 1 will not be on the MWC and will receive nothing. Subjects are not informed of this underlying payoff generating function; they are simply told that payoffs were determined based on a theory of decision-making in committees. Note, however, that we use binary choice problems in which the better committees for the subjects are the same regardless of whether we employ DPI or other power indices such as BzI (Banzhaf, 1965) and SSI (Shapley and Shubik, 1954).[3]

Problem	Choice 1	(Expected payoff)	Choice 2	(Expected payoff)
A	$[14; \mathbf{5}, 3, 7, 7]$	$(120 \times 2/3 \times 1/3)$	$[14; \mathbf{5}, 4, 6, 7]$	$(120 \times 3/4 \times 1/3)$
B	$[6; \mathbf{1}, 2, 3, 4]$	$(120 \times 1/3 \times 1/3)$	$[6; \mathbf{1}, 1, 4, 4]$	$(120 \times 2/3 \times 1/3)$
C	$[14; \mathbf{3}, 5, 6, 8]$	$(120 \times 2/3 \times 1/3)$	$[14; \mathbf{3}, 6, 6, 7]$	$(120 \times 3/4 \times 1/3)$
D	$[9; \mathbf{1}, 3, 5, 6]$	$(120 \times 1/3 \times 1/3)$	$[9; \mathbf{1}, 2, 6, 6]$	$(120 \times 2/3 \times 1/3)$

Table 1: Four binary choice problems used in the experiment. In each committee, subjects are all assigned to player 1, and the number of votes given to player 1 is shown in bold. Choice 2 generates a higher expected payoff for the subjects in all binary choice problems. In Problem A and Problem C, one option has two "large" voters who can form an MWC on their own, whereas the other option does not. In Problem B and Problem D, there is no such clear difference between the two options.

Note that in Problems A and C, one option has two large voters who can form an MWC on their own, whereas the other option does not. In Problems B and D, there is no such clear difference between the two options as there are two large voters who can form an MWC by themselves in both options. We hereafter refer to Problems A and C as "easy" binary choice problems and Problems B and D as "difficult" binary choice problem.

Subjects are faced with one of the following sequences of binary choice problems: A $\rightarrow$ B, B $\rightarrow$, C $\rightarrow$ D , or D $\rightarrow$ C (the order of problems is indicated by the arrows), where the first problem is used in the first 40 periods, and the second problem is used in the subsequent 20 periods. In Guerci et al. (2017), as shown in Table 1, the committee that generates a higher expected payoff for subjects (correct option) is Choice 2 for all problems. In the additional sessions, however, the correct option is Choice 1 in the first 40 periods and it is Choice 2 in the subsequent 20 periods. Subjects are not informed of what binary choice problems being given before those problems are shown on their monitors.

[3]Guerci et al. (2017) adopted DPI as their payoff generating function because, in all the experiments reported in Montero et al. (2008), Aleskerov et al. (2009), Esposito et al. (2012), Guerci et al. (2014), and Watanabe (2014), the most frequently observed winning coalitions were MWCs.

As noted at the beginning of this subsection, there are three treatments: (1) no feedback, (2) partial feedback, and (3) full feedback. For the no-feedback treatment, subjects are not informed of anything as the result of their committee choice until the session ends. For the partial-feedback treatment, each subject is informed of his or her own payoff in the committee he or she chooses. For the full-feedback treatment, after each choice subjects are informed of the payoffs of all four players in the committee they chose. There is a 30-second time limit for the choice stage and a 10-second limit for the feedback stage, regardless of the amount of feedback information. If a subject does not choose a committee within the 30 seconds of the choice stage, then he or she obtains zero points for that period. In this case, regardless of the treatment, in the feedback stage the subject receives a notice that he or she obtained nothing as a result of their failure to make a choice within the time limit.

If a subject makes an early choice, then a waiting screen is shown until all the subjects in the session have made their decisions. If all the subjects make their choice before the end of the 30-second time limit, then they all enter the feedback stage. For the no-feedback treatment, during the 10-second feedback stage subjects are shown a screen conveying the message "Please wait until the experiment continues." For the full-feedback and partial-feedback treatments, the relevant payoff information is displayed during these 10 seconds.

Cooper and Kagel (2003, 2008) found that discussion among subjects promotes meaningful learning in signaling games. In this experiment, any communication with the others is prohibited during each session. Subjects are also prohibited from note-taking, because it is considered as communication with themselves.

2.2 Subject's Monitor

The instructions for bandit experiments are intentionally simplified, because they examine whether subjects learn something from their experiences. For readers' convenience, we here illustrate the examples of what subjects see on their monitors in this experiment. In each session for the full-feedback treatment, subjects are provided a problem and other information on their monitors as follows.

Please choose one out of the following two committees (Choice 1 or Choice 2). Each committee decides a distribution of 120 points among four members. You are Member 1. In both committees, 22 votes are apportioned to those members and you have 5 votes. Any proposals of point distributions need 14 votes in favor to be adopted.

$$\text{Choice 1} \quad [14; \mathbf{5}, 3, 7, 7], \quad \text{Choice 2} \quad [14; \mathbf{5}, 4, 6, 7]$$

When subjects choose Choice 2 and MWC $(5, 6, 7)$ appears, they see, for instance, the following results on their monitor, regardless of any treatments.

You chose the following committee.

Choice 2: $[14; \mathbf{5}, 4, 6, 7]$.

Next, in the full-feedback treatment, subjects see

The committee decided to distribute 120 points this time as follows.
You obtained 40 points this time.

(**40**, 0, 40, 40)

on their monitors. In the partial-feedback treatment, the payoff distribution is not shown, but rather the following note is shown on their monitors:

You have obtained 40 points this time.

In the no-feedback treatment, the payoff distribution is not shown and simply

Please wait for a while.

is shown on the subjects' monitors.

2.3 Experimental Procedure

The sessions for Guerci et al. (2017) were conducted at the Institute of Social and Economic Research (ISER) at Osaka University in June 2014 and at the University of Tsukuba in November 2014. The subjects were undergraduate students recruited from all over the campus, but third- or fourth-year economics majors were excluded. Each session in Osaka (June) involved 20 subjects and each session in Tsukuba (Nov) involved 10 subjects, and thus 360 subjects participated in those sessions.

The sessions in Tsukuba (Nov) were conducted for a robustness check to the results obtained in the sessions in Osaka (June). The same results were obtained at two experimental sites, and thus the main results were reported in Guerci et al. (2017) with the pooled data. For the no-feedback treatment, the additional sessions were conducted also at the ISER in September 2014. Each session in Osaka (Sept) involved 10 subjects, and thus we had 40 subjects in total there. No one has ever participated in this experiment and every subject participated once in this experiment.

The experiment was computerized with z-Tree (Fischbacher, 2007). Each session lasted around 60 minutes including the time for administering the instructions and the post-experiment questionnaire. At the beginning of each session, subjects were provided with a written instruction upon arrival, and then the experimenter read it aloud. No communication among subjects was allowed. Subjects were allowed to ask questions regarding the instruction and they were given the answers which other subjects could hear. Thereafter, any information available to the subjects was provided through their computer screens.

At the end of each session, subjects were asked to write the reasoning behind their own choice in free format. For the payment scheme, we followed other bandit experiments (Meyer and Shi, 1995; Hu et al., 2013). In the instruction, each subject was informed that in addition to the show-up fee of JPY 1000, he or she would receive payment according to the total points he or she obtained over all 60 periods at a rate of 1 point = JPY 1. The average earning of our subjects was JPY 2500 (about 18 USD in 2014.)

3 Results

In Subsections 3.1 and 3.2, we reconfirm learning and meaningful learning of subjects who participated in sessions for the no-feedback treatment.

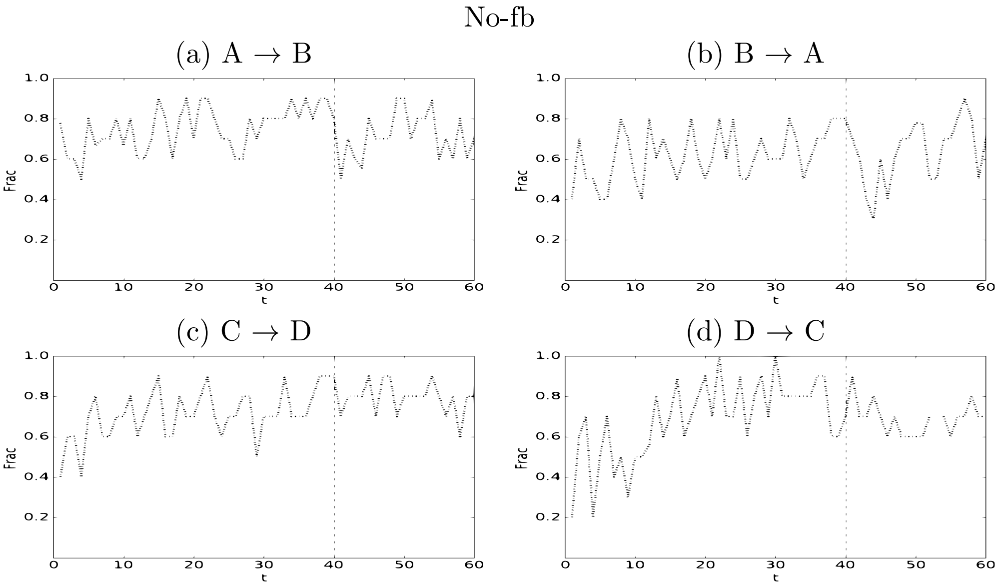

Figure 1: Time series of the percentage of subjects who chose the correct option among those who chose within the time limit in the additional sessions. 10 out of 40 subjects failed to make a choice within the time limit at least once in the 60 periods.

Figure 1 presents the time series of the percentage of subjects in additional sessions who chose the correct option among those who chose within the time limit in each of the four sequences; the dotted vertical line in each panel separates Period 40 and Period 41 to indicate that the problems were different in the periods before and after the line. In the additional sessions, as noted in Subsection 2.1, the correct option was Choice 1 in the first 40 periods, which was changed to Choice 2 in the subsequent 20 periods.

Recall that in Subsection 2.1, Problems A and C are named easy binary choice problems and Problems B and D are names difficult ones. In fact, at Period 1, as is seen in Figure 1, subjects found the correct option easier in Problems A and C than in Problems B and D, which was reported also in Guerci et al. (2017) for all the three treatments. For each binary choice problem, we could not reject the null hypothesis that the percentage of subjects who chose the correct option among those who chose within the time limit at Period 1 was equal between the additional sessions and the sessions conducted for Guerci et al. (2017); the p-values for the two-sided χ^2 test are 0.526, 0.559, 0.853, and 0.673 in Problems A, B, C, and D, respectively. Thus, it is not inappropriate to compare the additional data with those used in Guerci et al. (2017) in the following subsections. The features of the additional sessions are summarized in Table 2. There was no large difference in the average amounts of payments among Osaka (June), Osaka (Sept), and Tsukuba (Nov).

In what follows, no-feedback, partial feedback, and full-feedback treatments are abbreviated as No-fb, Part-fb, and Full-fb, respectively. We sometimes indicate the number of subjects who participated in sessions: No-fb (10) for Osaka (Sept), No-fb(30)), Part-fb(30), and Full-fb (30) for Osaka (June) and Tsukuba (Nov).

Table 2: Features of the additional sessions.

session no.	sequence of binary choice	show-up fee (JPY)	point-to-JPY ratio	# of subj.	session date	avg. payment to subject
1	A →B	1000	1.0	10	Sept.24, 2014	2552
2	B →A	1000	1.0	10	Sept.24, 2014	2442
3	C →D	1000	1.0	10	Sept.25, 2014	2628
4	D →C	1000	1.0	10	Sept.25, 2014	2472

3.1 Learning the Correct Option in Additional Sessions

In Subsections 3.1 and 3.2, we analyze the data in additional sessions. Let FR_k^i denote the relative frequency of periods in which subject i chose the correct option within the k-th block of 5 consecutive periods, that is, from $5(k-1)+1$ to $5k$. FR_2^i is, e.g., the number of times subject i chose the correct option from Period 6 to Period 10, divided by 5. The change in the relative frequencies that subject i chose the correct option between the l-th block and the m-th block is defined as

$$\Delta\mathrm{FR}_{l,m}^i = \mathrm{FR}_l^i - \mathrm{FR}_m^i.$$

Let $\Delta\mathrm{FR}_{l,m}$ denote a vector whose i-th component is $\Delta\mathrm{FR}_{l,m}^i$. For each of 4 sequences of binary choice problems, we had 10 subjects in Osaka (Sept) in which No-fb was conducted, and thus we have 10 observations of $\Delta\mathrm{FR}_{l,m}$.

We expect experience in introspective thinking to improve learning for choosing the correct option even without any payoff-related feedback information. For each binary choice problem, the p-value for the one-tailed signed-rank (SR) test for each feedback treatment is thus reported in Table 3, where the null hypothesis is $\Delta\mathrm{FR}_{2,1} \leq 0$ ($\Delta\mathrm{FR}_{8,1} \leq 0$) and the alternative hypothesis is $\Delta\mathrm{FR}_{2,1} > 0$ ($\Delta\mathrm{FR}_{8,1} > 0$).

	Problem A	Problem B	Problem C	Problem D
$\Delta\mathrm{FR}_{2,1}$	0.2948	0.1763	0.1990	0.3392
$\Delta\mathrm{FR}_{8,1}$	**0.0339**	**0.0463**	**0.0125**	**0.0142**

Table 3: P-values of one-tailed SR test in the sessions for the no-feedback treatment.

As shown in Table 3, for all problems, values in $\Delta\mathrm{FR}_{2,1}$ were not significantly greater than zero but values in $\Delta\mathrm{FR}_{8,1}$ were significantly greater than zero. Thus, we see that, overall, the distributions of $\Delta\mathrm{FR}_{8,1}$ lie towards the right of those for $\Delta\mathrm{FR}_{2,1}$.

Observation 1. *It was reconfirmed for all binary choice problems that subjects learned to choose the option with higher expected payoffs even without any feedback information regarding their payoffs.*

Guerci et al. (2017) reported that in Problems B and D, subjects learned to choose the option with higher expected payoffs even without any payoff-related feedback information. In the next subsection, Problems B and D are thus the candidates of the binary choice problems in which subjects meaningfully learned the underlying feature of weighted voting.

3.2 Meaningful Learning (Learning Transfer) in Additional Sessions

We say that subjects are experienced if they have completed their choice in the first 40 periods. For each binary choice problem, we say that meaningful learning was observed if the following criteria (a) and (b) were both satisfied; (a) There was a significant increase in number of subjects who chose the correct option between Period 1 and Period 41. (b) There was a significant increase from FR_1 to FR_9. Note that in criteria (a) and (b), subjects who chose at Period 1 or for FR_1 and those at Period 41 or for FR_9 are faced with different sequences of binary choice problems.

	No-fb (10)	No-fb (30)	Part-fb (30)	Full-fb (30)
Problem A				
Inexperienced	7	20	25	22
Experienced	7	15	16	17
p-value (χ^2)	1.000	0.190	**0.012**	0.176
Problem B				
Inexperienced	4	9	13	9
Experienced	5	21	8	10
p-value (χ^2)	0.653	**0.002**	0.176	0.781
Problem C				
Inexperienced	4	19	22	19
Experienced	9	23	22	19
p-value (χ^2)	**0.019**	0.260	1.000	1.000
Problem D				
Inexperienced	2	8	5	9
Experienced	7	21	16	15
p-value (χ^2)	**0.025**	**<0.001**	**0.003**	0.114

Table 4: Numbers of subjects who chose the correct option at Period 1 and Period 41 for the binary choice problems used in (No-fb (10) and No-fb(30), Part-fb(30)), and Full-fb (30). For each treatment, the number of subjects who participated in sessions for the treatment is noted in parentheses. The p-values for χ^2 tests are reported for comparison between inexperienced and experienced subjects.

First, we consider whether criterion (a) is satisfied. The left-most column No-fb (10) in Table 4 presents the numbers of subjects who chose the correct option at Period 1 and Period 41 in each binary choice problem in the additional sessions for the no-feedback treatment. Table 4 also lists those numbers reported in Guerci et al. (2017) in the columns No-fb (30), Part-fb (30), and Full-fb (30) for reference.

The p-values for the χ^2 test are also reported in Table 4, where the null hypothesis is that the percentages of inexperienced subjects and experienced subjects who chose the correct options are the same when they first encountered the same problem. Significantly more experienced subjects chose the correct options than inexperienced subjects in No-fb (30) for Problem B ($p = 0.002$), in No-fb (10)) for Problem C ($p = 0.019$), and in No-fb (10), No-fb (30), Part-fb (30) for Problem D ($p = 0.025$, $p < 0.001$, and $p = 0.003$, respectively). Thus, those five cases satisfied criterion (a).

	No-fb (10)	No-fb (30)	Part-fb (30)	Full-fb (30)
Problem A				
p-value (perm, MW)	0.445	0.867	0.854	0.547
p-value (KS)		0.913	0.889	0.283
Problem B				
p-value (perm, MW)	0.376	**0.049**	0.336	0.538
p-value (KS)		**0.030**	0.299	0.297
Problem C				
p-value (perm, MW)	0.107	0.052	0.453	0.994
p-value (KS)		**0.009**	0.536	0.297
Problem D				
p-value (perm, MW)	**0.001**	<**0.001**	0.264	0.791
p-value (KS)		<**0.001**	0.103	0.990

Table 5: P-values for testing $FR_1 = FR_9$. Guerci et al. (2017) reported the p-values for the Mann-Whitney U-test (MW) and the Kolmogorov-Smirnov test (KS) for No-fb (30), part-fb (30), and Full-fb (30). For No-fb (10), the p-values are not calculated with the MW test and the KS test but with the permutation test (perm) due to the small sample size. In No-fb (10), criteria (a) and (b) are satisfied only for Problem D.

Next, we consider whether criterion (b) is satisfied by comparing the relative frequency in which inexperienced subject i chose the correct option within the first block of 5 consecutive periods (Periods 1 to 5) with the relative frequency in which experienced subject j chose the correct option within the first block of 5 consecutive periods when they faced the same problem within the ninth block of 5 consecutive periods (Periods 41 to 45). Table 5 lists p-values for the Mann-Whitney U-test (MW) test as well as the Kolmogorov-Smirnov test (KS) of the distributions of FR_1 (inexperienced subjects) and FR_9 (experienced subjects) for each feedback treatment, where the null hypothesis is $FR_1 = FR_9$.

In Part-fb (30) and Full-fb (30), the p-values for the MW and KS tests in Table 5 suggest that there be no significant difference between FR_1 and FR_9 for all problems. Thus, meaningful learning was not observed for the partial-feedback and full-feedback treatments. In sessions for No-fb (30), however, there were significant differences between FR_1 and FR_9 in Problem B ($p = 0.049$ for MW and 0.030 for KS, respectively) and in Problem D ($p < 0.001$ for MW and < 0.001 for KS, respectively). For Problems B and D, criteria (a) and (b) were thus both satisfied. Therefore, Guerci et al. (2017) reported that meaningful learning was observed only for the no-feedback treatment in Problems B and D.

In No-fb (10), there was a significant difference between FR_1 and FR_9 in Problem D ($p = 0.001$ for perm), whereas there was no such a difference in Problem B ($p = 0.376$ for perm). For Problem C, criterion (a) was truly satisfied, as was confirmed above but criterion (b) was not. Thus, only for Problem D, criteria (a) and (b) are both satisfied.

Thus, we have the following observation. This observation was not made by chance. For Problem D in No-fb (10), the percentage of subjects who chose the correct option was 70-90% in B9 in sequence C $\rightarrow$ D (Figure 1). In the post experimental questionnaire, at least 6 subjects who participated in sequence C $\rightarrow$ D noted the proper reasoning for their choice (Appendix B).

Observation 2. *Statistically significant evidence of meaningful learning was reconfirmed in Problem D in the additional sessions for the no-feedback treatment.*

Guerci et al. (2017) reported that meaningful learning was observed in Problems B and D, but in additional sessions it was observed solely in Problem D. Recall that at Period 1 in the additional sessions, the percentage of subjects who chose the correct option in Problems D was lowest (Figure 1). Under criteria (a) and (b), we might be able to detect meaningful learning more easily in Problem D than in Problem B. These findings are summarized in the following statement.

Result 1. *Hypothesis 1 was affirmatively confirmed in a sequence that starts with an easy problem and then gives a difficult one.*

According to Observation 2, we have the following result. We conclude that meaningful learning without any payoff-related feedback information is not the consequence of the inertia on subjects' choice, because the correct option was Choice 1 in the first 40 periods and it was changed to Choice 2 in the subsequent 20 periods.

4 Subjects' Search Behavior

Meaningful learning was observed only in No-fb, according to Observation 2 and Guerci et al. (2017). Then, what hindered subjects from meaningfully learning the underlying structure of weighted voting, when immediate feedback information was given to them? For each binary choice problem, this section compares the behavior of subjects who were searching for the correct options among treatments. Prior to the comparison, we refer to the length of thinking time subjects spent for their decision, because we do not have to care about the effect of thinking time on their behavior in the comparison if they spent the same length of time for choosing options among sessions for all treatments.

4.1 Length of Thinking Time

We recorded the remaining time in seconds until the time limit (30 seconds); for each subject, his or her thinking time is computed by subtracting the remaining time from the time limit. Table 6 presents the average lengths of remaining time measured in seconds for subjects to make their choice until the time limit in the 1st block and the 9th block of 5 consecutive periods in sessions for each treatment. The p-values for the Kruskal-Wallis test are provided in the rightmost column, where the null hypothesis is that the lengths of remaining time for subjects to choose in the same 5 consecutive periods when they were first faced with a binary choice problem are on average the same among all treatments. In Problems C and D, there was no significant difference in average length of remaining time experienced subjects had among all treatments.

Table 7 presents p-values for the one-sided Steel test for multiple comparison of average lengths of remaining time subjects had in 1st block and 9th block of 5 consecutive periods. The alternative hypothesis is that as compared to No-fb (10) (control group), they had a longer remaining time on average in No-fb (30), Part-fb (30), and Full (30) (treatment groups), respectively. For every pair of the control group and the treatment group, no significant difference was observed in average length of remaining time experienced subjects had in Problems C and D. Thus, we note this point as a remark.

Remark: We found that there was no significant difference in length of thinking time the experienced subjects who were faced with Problem D spent among treatments.

As stated in Observation 2, meaningful learning was reconfirmed in Problem D in No-fb. In the same length of thinking time, what occurred in subjects' behavior when they were given feedback information? We note a comment on this point in subsection 4.3.

	No-fb (10)	No-fb (30)	Part-fb (30)	Full-fb (30)	p-value (KW)
Problem A					
Inexperienced	14.260	15.380	20.547	17.380	**0.003**
Experienced	18.200	15.533	20.102	22.287	<**0.001**
p-value (BM)	0.086	0.812	0.816	<**0.001**	
Problem B					
Inexperienced	12.060	13.980	16.753	18.113	**0.004**
Experienced	17.700	17.120	22.567	23.373	<**0.001**
p-value (BM)	**0.002**	**0.039**	<**0.001**	<**0.001**	
Problem C					
Inexperienced	14.600	12.247	17.907	16.213	<**0.001**
Experienced	20.760	19.149	21.353	21.167	0.240
p-value (BM)	**0.049**	<**0.001**	**0.004**	<**0.001**	
Problem D					
Inexperienced	14.240	13.400	15.847	17.087	**0.046**
Experienced	19.800	18.913	20.800	20.667	0.195
p-value (BM)	**0.016**	<**0.001**	<**0.001**	**0.022**	

Table 6: Average lengths of remaining time in seconds for subjects to choose in the 1st block (for inexperienced subjects) and the 9th block (for experienced subjects) of 5 consecutive periods. The p-values for the two-sided Brunner-Munzel test (BM) are provided for the comparison of average lengths of remaining time between inexperienced subjects and experienced subjects who were faced with the same binary choice problem. The p-values for the Kruskal-Wallis test (KW) are also provided for testing whether there is no difference in average lengths of remaining time subjects had in the same 5 consecutive periods among all treatments.

4.2 Switching Options

The statistical test results shown in Subsection 3.2 imply that meaningful learning was not observed in sessions for both the partial-feedback and the full-feedback treatments. Then, what hindered subjects from meaningfully learning in sessions for those treatments with feedback information? As noted in Guerci et al. (2017), in the post-experimental questionnaire, some subjects in sessions for the partial-feedback and full-feedback treatments reported that they changed their choices when they received zero points, while they did not when they received a non-zero amount of points. If subjects had not yet been convinced of what they had learned, then it would be plausible for them to take this type of "win-stay-lose-shift (WSLS)" strategy (Nowak and Sigmund, 1993).

		No-fb (30)	Part-fb (30)	Full-fb (30)
Problem A				
Inexperienced	No-fb (10)	0.402	**0.008**	0.066
Experienced	No-fb (10)	0.968	0.238	0.034
Problem B				
Inexperienced	No-fb (10)	0.480	**0.030**	**0.008**
Experienced	No-fb (10)	0.604	**0.004**	**0.004**
Problem C				
Inexperienced	No-fb (10)	0.946	0.204	0.427
Experienced	No-fb (10)	0.916	0.620	0.315
Problem D				
Inexperienced	No-fb (10)	0.890	0.383	0.223
Experienced	No-fb (10)	0.661	0.078	0.287

Table 7: P-values for the one-sided Steel test for multiple comparison of average lengths of remaining time for subjects to choose that were observed in 1st block and 9th block of 5 consecutive periods, where the alternative hypothesis (H_1) is that as compared to the control group No-fb (10), they had a longer remaining time on average in the treatment group No-fb (30), Part-fb (30), and Full-fb (30), respectively.

In this experiment, subjects received 0 points or 40 points, as shown in Table 12 in Subsection 4.3), in all binary choice problems. Tables 8 - 11 list the frequencies of observing 0 points and 40 points (*freq*), the frequencies of switching choices immediately after observing 0 points and 40 points (*switch*), the ratios of *switch* to *freq* (*ratio*), for the partial-feedback treatment and the full-feedback treatment in the 1st (B1), 8th (B8), 9th (B9), and 12th (B12) blocks of 5 consecutive periods, respectively.

The p-value for the two-sided Fisher exact test is reported for each binary choice problem, where the null hypothesis is that switching choices immediately after observing 0 points and switching choices immediately after observing 40 points were equally likely to be observed. In the following analysis, we confine attention to Problem D in sequence C $\rightarrow$ D, because meaningful learning was observed in No-fb (30) as well as in No-fb (30) for Problem B but it is observed only in No-fb (10) for Problem D (Observation 2).

Table 8: Frequency of switching choices in B1: two-sided Fisher test

Part-fb (30)				Full-fb (30)			
		0 point	40 points			0 point	40 points
Problem A	freq	43	77	Problem A	freq	44	76
	switch	23	18		switch	25	19
	ratio	0.535	0.234		ratio	0.568	0.250
	p-value		**0.001**		p-value		**0.001**
B	freq	53	66	B	freq	65	55
	switch	30	15		switch	36	13
	ratio	0.566	0.227		ratio	0.554	0.236
	p-value		<**0.001**		p-value		**0.001**
C	freq	27	90	C	freq	39	81
	switch	12	30		switch	20	28
	ratio	0.444	0.333		ratio	0.513	0.346
	p-value		0.361		p-value		0.111
D	freq	57	62	D	freq	62	57
	switch	37	21		switch	37	13
	ratio	0.649	0.339		ratio	0.597	0.228
	p-value		**0.001**		p-value		<**0.001**

Table 9: Frequency of switching choices in B8: two-sided Fisher test

Part-fb (30)				Full-fb (30)			
		0 point	40 points			0 point	40 points
Problem A	freq	45	105	Problem A	freq	49	101
	switch	15	23		switch	15	22
	ratio	0.333	0.219		ratio	0.306	0.218
	p-value		0.155		p-value		0.313
B	freq	59	91	B	freq	72	78
	switch	20	10		switch	21	9
	ratio	0.339	0.120		ratio	0.292	0.115
	p-value		**0.001**		p-value		**0.008**
C	freq	43	107	C	freq	39	111
	switch	10	31		switch	12	18
	ratio	0.233	0.290		ratio	0.282	0.162
	p-value		0.547		p-value		0.063
D	freq	52	97	D	freq	62	57
	switch	15	13		switch	37	13
	ratio	0.289	0.134		ratio	0.452	0.210
	p-value		**0.028**		p-value		<**0.001**

Table 10: Frequency of switching choices in B9: two-sided Fisher test

Part-fb (30)				Full-fb (30)			
		0 point	40 points			0 point	40 points
Problem A	freq	28	92	Problem A	freq	40	80
	switch	15	13		switch	12	12
	ratio	0.536	0.130		ratio	0.300	0.150
	p-value		<**0.001**		p-value		0.088
B	freq	58	62	B	freq	55	65
	switch	23	7		switch	28	14
	ratio	0.397	0.113		ratio	0.509	0.215
	p-value		<**0.001**		p-value		**0.001**
C	freq	36	83	C	freq	35	85
	switch	16	16		switch	20	18
	ratio	0.444	0.193		ratio	0.571	0.212
	p-value		**0.007**		p-value		<**0.001**
D	freq	54	66	D	freq	52	68
	switch	20	14		switch	31	12
	ratio	0.370	0.212		ratio	0.566	0.177
	p-value		0.068		p-value		<**0.001**

Table 11: Frequency of switching choices in B12: two-sided Fisher test

Part-fb (30)				Full-fb (30)			
		0 point	40 points			0 point	40 points
Problem A	freq	60	85	Problem A	freq	63	87
	switch	23	11		switch	19	9
	ratio	0.383	0.129		ratio	0.302	0.104
	p-value		<**0.001**		p-value		**0.003**
B	freq	36	114	B	freq	48	102
	switch	4	15		switch	10	12
	ratio	0.111	0.132		ratio	0.208	0.118
	p-value		>0.999		p-value		0.215
C	freq	56	93	C	freq	54	95
	switch	24	17		switch	18	11
	ratio	0.429	0.183		ratio	0.333	0.116
	p-value		**0.002**		p-value		**0.002**
D	freq	41	108	D	freq	45	105
	switch	9	25		switch	22	25
	ratio	0.220	0.232		ratio	0.489	0.238
	p-value		>0.999		p-value		**0.004**

In sequence C $\to$ D in Part-fb (30), it is not plausible that in B1 and B8 inexperienced subjects took the WSLS strategy in Problem C ($p = 0.361$ and $p = 0.290$) and also it is not plausible that subjects who experienced Problem C from B1 to B8 and encountered Problem D in B9 took the WSLS strategy ($p = 0.068$). Moreover, the null hypothesis was not rejected also in B12 ($p > 0.999$).

Observation 3. *In the partial-feedback treatment, subjects did not take the win-stay-lose-shift strategy in Problem C and subjects who experienced Problem C also did not take that strategy in Problem D.*

It is not plausible that inexperienced subjects learned something in Problem C ($p = 0.768$ for the one-sided SR test for the null hypothesis $\Delta FR_{8,1} \leq 0$) according to Figure 3 (c) in Guerci et al. (2017). As noted above, however, it is also not plausible that subjects who experienced Problem C took the WSLS strategy in Problem D. What search behavior did they take? The answer is shown in subsection 4.3.

In sequence C $\to$ D in Full-fb (30), it is not plausible that in B1 subjects took the WSLS strategy in Problem C ($p = 0.111$) and yet they did in B8 ($p = 0.063$), as in the case of Part-fb (30). It is, however, inferred that subjects who experienced Problem C from B1 to B8 and encountered Problem D took the WSLS strategy ($p < 0.001$) in B9 and that they still took the WSLS strategy ($p = 0.004$) in B12. We have the following observation.

Observation 4. *In the full-feedback treatment, inexperienced did not take the win-stay-lose-shift strategy in Problem C but subjects who experienced Problem C took the win-stay-lose-shift strategy in Problem D.*

Let us look at another sequence. It is inferred that in B9 subjects who experienced Problem A from B1 to B8 and encountered Problem B took the WSLS strategy ($p < 0.001$), where "lose-shift" was chosen at about 40% (=0.397) after 0 points realized while "win-stay" was chosen at about 89% (=1-0.113) after 40 points realized, although the null hypothesis was rejected in B12. It is inferred that subjects who experienced Problem A from B1 to B8 and encountered Problem B took the WSLS strategy ($p = 0.001$) in B9, although the null hypothesis was rejected in B12. All test results are summarized in Table 13 in subsection 4.3 together with the test results of random choice of runs in each sequence of the options.

We summarize these test results including the ones stated in Observations 3 and 4 in the following result; there were some sequences in which subjects took the win-stay-lose-shift strategy, when feedback information is provided immediately after their choice.

Result 2. *Hypothesis 2 was partly verified.*

4.3 Random Choice of Runs

In reference to Observation 3, we could not conclude that in sequence C $\rightarrow$ D, the WSLS strategy hindered subjects from meaningful learning in sessions for the partial-feedback treatment. In this subsection, we consider whether subjects randomly chose the runs of options in their search for the correct ones.

Problem A	Choice 1	[14; 5, 3, 7, 7]	Choice 2	[14; 5, 4, 6, 7]
	$(7_1, 7_2)$	(0, 0, 60, 60)	(5, 4, 6)	(40, 40, 40, 0)
	$(5, 3, 7_1)$	(40, 40, 40, 0)	(5, 4, 7)	(40, 40, 0, 40)
	$(5, 3, 7_2)$	(40, 40, 0, 40)	(5, 6, 7)	(40, 0, 40, 40)
			(4, 6, 7)	(0, 40, 40, 40)
Problem B	Choice 1	[6; 1, 2, 3, 4]	Choice 2	[6; 1, 1, 4, 4]
	(2, 4)	(0, 60, 0, 60)	$(4_1, 4_2)$	(0, 0, 60, 60)
	(3, 4)	(0, 0, 60, 60)	$(1_1, 1_2, 4_1)$	(40, 40, 40, 0)
	(1, 2, 3)	(40 40, 40, 0)	$(1_1, 1_2, 4_2)$	(40, 40, 0, 40)
Problem C	Choice 1	[14; 3, 5, 6, 8]	Choice 2	[14; 3, 6, 6, 7]
	(6, 8)	(0, 0, 60, 60)	$(3, 6_1, 6_2)$	(40, 40, 40, 0)
	(3, 5, 6)	(40, 40, 40, 0)	$(3, 2, 6_1)$	(40, 40, 0, 40)
	(3, 5, 8)	(40, 40, 0, 40)	$(3, 2, 6_2)$	(40, 0, 40, 40)
			$(6_1, 6_2, 7)$	(0, 40, 40, 40)
Problem D	Choice 1	[9; 1, 3, 5, 6]	Choice 2	[9; 1, 2, 6, 6]
	(3, 6)	(0, 60, 0, 60)	$(6_1, 6_2)$	(0,0, 60, 60)
	(5, 6)	(0, 0, 60, 60)	$(1, 2, 6_1)$	(40, 40, 40, 0)
	(1, 3, 5)	(40, 40, 40, 0)	$(1, 2, 6_2)$	(40, 40, 0, 40)

Table 12: MWCs and payoff vectors. The MWCs in $[14; 5, 3, 7, 7]$ are, e.g., written here as $(5, 3, 7_1)$, $(5, 3, 7_2)$, and $(7_1, 7_2)$ by the votes apportioned to the members, not with references to the specific players.

Under the null hypothesis, the number of runs of the options each subject chooses is a random variable. When the payoffs associated with options (shown in Table 12) are stochastically determined, however, the null hypothesis is not rejected as often in the runs test, even if the sequence of options is generated by the WSLS strategy.[4] We thus prioritize the test result for the WLSL strategy over the result in the runs test, when it is inferred that subjects took the WSLS strategy in both B1 and B8 or in both B9 and B12.

[4]In Table 12, denote by 1-1, 1-2, and 1-3 payoff vectors (0, 60, 0, 60), (0, 0, 60, 60), and (40, 40, 40, 0), respectively when Choice 1 is chosen in Problem D and by 2-1, 2-2, and 2-3 (0, 0, 60, 60), (40, 40, 40, 0), and (40, 40, 0, 40), respectively when Choice 2 is chosen. Assume that when Choice k $(= 1, 2)$ is successively chosen, payoff vectors realize in the order of k-1$\rightarrow$ k-2 $\rightarrow$ k-3 $\rightarrow$ k-1 $\rightarrow \cdots$, and assume in addition that when the alternative option is once chosen after observing payoff vector k-i (i=1, 2, 3) and then Choice k is chosen again, the order resumes from the payoff vector next to k-i. Given that Choice 1 is chosen at Period 1, the WSLS strategy generates the following sequence of 20 choices: $1, 2, 1, 2, 2, 2, 1, 1, 2, 2, 2, 1, 2, 2, 2, 1, 1, 2, \ldots$. The p-value for the runs test applied to the sequence of those choices is 0.8391, and thus the null hypothesis was not rejected, although they were generated systematically with the WSLS strategy.

Table 13 presents the numbers of subjects each of whom is counted if the null hypothesis was rejected in the runs test for each binary choice problem in sessions for each treatment. The sample size of 5 consecutive periods would be too small to conduct the runs test. The runs test was thus applied to a sequence of the options each identical subject chose in a binary choice problem in Periods 21-40 and in another one in Periods 41-60, respectively. In the table, "W" is put when the WSLS strategy was observed, while "w" was used when it was observed only in the first 5 consecutive periods in each binary choice problem.

	No-fb (10)	No-fb (30)	Part-fb (30)		Full-fb (30)	
Problem A						
Inexperienced	5	22	12		11	
Experienced	3	15	9	W	15	
Problem B						
Inexperienced	4	22	13	W	11	W
Experienced	5	13	8	w and rand	8	w and rand
Problem C						
Inexperienced	5	14	9	**rand**	12	
Experienced	4	18	10	W	13	W
Problem D						
Inexperienced	7	21	12	W	7	W
Experienced	5	15	7	**rand**	7	**W**

Table 13: Numbers of subjects each of whom is counted if the null hypothesis (H_0) was rejected, where H_0 is that the number of runs in a sequence of the options each subject chose in Periods 21-40 (in Periods 41-60) is a random variable. "W" indicates that we can infer that the WSLS strategy was taken in both B1 and B8 for the inexperienced subjects and in both B9 and B12 for the experienced subjects, respectively, and "w" was used when the WSLS strategy was observed only B9. Another mark "rand" indicates that subjects did not search the correct options with any certain rules, according to the test result for the random choice of runs. Letters in bold indicates the test results in sequence C $\rightarrow$ D.

As shown in Subsection 3.1, for all four binary choice problems we examined, it was reconfirmed that subjects learned to choose the correct options in sessions for No-fb (10) (Observation 1). Table 13 shows that in Problem B, 40% of inexperienced subjects searched for the correct options with a certain rule in sessions for No-fb (10). Thus, we presume in the following analysis that subjects did not search for the correct options with any certain rules, if at most 30% of subjects are counted for each of whom the null hypothesis was rejected in the runs test and if they did not take the WSLS strategy. In Table 13, "rand" is put when subjects did not search for the correct options with any certain rules.

We found in subsection 4.1 that there was no significant difference in length of thinking time the experienced subjects spent in Problem D among treatments (Remark). In almost the same length of thinking time, how did those subjects search for the correct options, when they were given feedback information? Observation 3 states that in Part-fb (30) subjects who experienced Problem C did not take the WSLS strategy in Problem D. Table 13 shows, however, that there were only 7 experienced subjects in Problem D are counted for each of whom the null hypothesis was rejected in the runs test. Observation 4 states that in Full-fb (30) subjects who experienced Problem C did take the win-stay-lose-shift strategy in Problem D. We have the following observation, according the 30% criterion noted above and the priority of the test result for the WLSL strategy.

Observation 5. *In the partial-feedback treatment, subjects did not search for the correct options with any certain rules in Problem D. In the full-feedback treatment, experienced subject took the win-stay-lose-shift strategy in Problem D.*

In Problem D, experienced subjects chose their options in almost the same thinking time, regardless of the treatments in which they participated. Observation 5 thus more neutrally implies the reason why subjects could not meaningfully learn the underlying relationship between payoff distributions and vote apportionments, when payoff-related feedback information was provided immediately after their choice.

Finally, let us consider the cases of Problem B, because meaningful learning was observed also in No-fb (30). What occurred in Part-fb (30) and Full-fb (30)? In each of those feedback treatments, as shown in Table 13, experienced subjects did take the WSLS strategy in B9, and there were only 8 experienced subjects who were counted for each of whom the null hypothesis was rejected in the runs test. Thus, we have the following observation.

Observation 6. *In both partial-feedback treatment and the full-feedback treatment, experienced subjects took the win-stay-lose-shift strategy in some consecutive periods after they were first faced with Problem B but finally they did not search for the correct options with a certain rule.*

Table 13 indicates that when subjects received immediate feedback information, there are some cases where we could not reject the null hypothesis that that they did not search for the correct options with a certain rule. We summarize those findings including Observations 5 and 6 as the following result. Immediate payoff-related feedback information sometimes induced subjects to randomly choose the runs of options.

Result 3. *Hypothesis 3 was partly verified.*

5 Final Remarks

In this section, we note brief remarks on the improvement of our experimental design and possible questions for future research, based on our results and observations. One is the verification of implications derived by theoretical models, and the other is the relationship between subjects' cognitive ability and their meaningful learning.

Potential Theoretical Models

Results 2 and 3 might suggest that subjects' inference on the relationship between nominal voting weights and actual payoffs be mistaken or confused when feedback information was given immediately after their choice. This idea is still unverified, but if it is true, then we can say that in the sessions for no-feedback treatment, subjects might deeply infer the underlying structure of weighted voting without such misleading or confusing information by being kept away from the feedback information. Even in sessions for full-feedback treatments, however, there were some subjects who could succeed in meaningful learning.

What feedback information then induced those subjects to meaningfully learn the underlying structure of weighted voting meaningfully? In this experiment, subjects were not given the information about cumulative payoffs but provided with instantaneous payoffs given immediately after their decision, and they were prohibited from taking any notes during the sessions. By this lack of sufficient memory on the outcomes that were realized by their previous choice, immediate payoff-related feedback information might confuse subjects' inference on the underlying structure of weighted voting. From the viewpoint of the effect of subjects' memory on their learning, case-based decision theory (CBDT) might also provide theoretical reasoning for some patterns of the search behavior observed in the partial-feedback and full-feedback treatments. (Gilboa and Schmeidler (2001) is a comprehensive review of the CBDT written by the founders of the theory.)

The CBDT describes how people make an analogy of the circumstances surrounding them by the past experiences when they are ignorant of the structure. If we assumed that subjects were memoryless, then the CBDT would provide a possible explanation of the WSLS strategy of subjects who gave up learning. On the other hand, if subjects in our experiment could take notes on realized payoff distributions, then they might infer the correct options. At the beginning of the experimental instruction, however, subjects were informed that they were prohibited from taking notes during the session. Thus, the information on the cumulative payoffs they earned might assist subjects in inferring the expected payoffs of the options they had chosen. We cannot verify what memory subjects have during the sessions, but we can record what information subjects see. Therefore, a "mouse-tracking" experiment is an alternative way for capturing the effect of subjects' memory on their leaning behavior.[5]

There is another mathematical model to test the theoretical implications behind our observations. Grant et al. (2017), for example, proposed a model of learning in which complete lack of information regarding the underlying data generating process is expressed as a (maximal) family of priors. In the replication studies of human learning, Arifovic et al. (2006) showed that standard models fail to replicate human behavior in a repeated game of battle of the sexes, and Erev et al. (2010) reported that those models do not perform well in predicting how people behave in market entry games. It would thus be valuable to detect the implications derived from these theoretical models.

Subjects' Cognitive Ability

Albeit on a trial basis, subjects' cognitive ability scores were measured by the Raven's APM test in the sessions for Guerci et al. (2017) and the additional sessions conducted by the author. The Raven's test is a well-known test that measures subject's ability for visual pattern recognition, and there are three versions; Colored Progressive Matrices (CPM), Standard Progressive Matrices (SPM), and Advanced Progressive Matrices (APM), in ascending order of difficulty. The Raven's APM test is composed of 48 questions in total, but in those sessions 16 questions were selected so that that the subjects could answer those questions within 10 minutes after the bandit experiment.[6] The average score was about 12.3 for subjects who participated in the sessions for Guerci et al. (2017) and it was 12.1 for those who participated in the additional sessions conducted by the author.

Ogawa et al. (2020) conducted the same experiment as the one designed by Guerci et al. (2017) at four experimental sites the subjects' characteristics of which were different, and they reported that meaningful learning was observed only at an experimental site where subjects had, on average, higher scores (11.6) of the Raven's APM test than those had at other sites. Accordingly, it would be inferred that meaningful learning could be observed at the experimental sites where the subjects' cognitive ability scores were relatively high. This feature should be reconfirmed by designing a more comprehensive experiment.

[5]In the mouse-tracking experiment, votes and payoffs of the committee members were "hidden" from subjects in windows on their monitors; using a computer mouse, he or she needed to bring the cursor to the windows and click on them to view the hidden information. Almost all other aspects are the same as those for the full-feedback treatment in our experiment.

[6]The question numbers are 1, 4, 7, and 10 from Set I and questions 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, and 34 from Set II.

Declarations

Conflict of Interest

The author declares that there is no conflict of interest. Eric Guerci and Nobuyuki Hanaki gave their permission to reuse the data to the author. Their permission emails can be shown to the editors of this journal.

Availability of Data and Material

All raw and processed data as well as zTree codes are available upon requests.

Code Availability

All data were processed with *Excel Statistics*. The software information is available at the following website, although the contents are written in Japanese.

`https://bellcurve.jp/ex/`

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committees and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Appendix A: Instructions

The instructions for bandit experiments are in general less informative to subjects. We attach the instructions here for showing to the readers that they were actually simple.

Instructions

Welcome! Thank you for participating in this experiment today. You will be paid 1000 JPY for your participant and an additional reward that ranges from 0 to 3200 JPY depending on your choice and performance in the experiment. At first,

- Please follow the instructions of the experimenter.
- Please do not take notes during this session.
- Please remain quiet and especially do not talk with other participants.
- Please do not look at what other participants are doing.
- During the experiment, please maintain an upright posture without leaning on the backrest.
- Do absolutely nothing other than the operation that you are instructed to do.
- Please turn off your mobile phone and definitely refrain from using it.
- If you have any questions or require assistance, please silently raise your hand.

You will be asked to repeatedly make a simple choice between two options. Imagine that you need to represent your interests within a voting committee. This committee decides how to divide 120 points among its members. The committee has three other members, and each member has a predetermined number of votes, which may be different from one to the other. The committee will make a decision only when a proposal receives the predetermined required number of votes. You will be told what is the required number of votes. If more than one proposal is put before the committee, the members cannot vote for multiple proposals by dividing their allocated number of votes. A member can vote for only one proposal, and all of his/her votes must be cast for that proposal.

You are asked to choose which of the two possible committees you prefer to join. You will be informed of the number of votes allocated to each of the four members of the committee (including you), and the number of votes required for a proposal to be approved. The number of votes you have will always be indicated with the label YOU.

full-feedback treatment

There are a total of 60 periods. At each period, you have 30 seconds to make your choice between the two committees. If you do not make a choice within the 30 seconds at one period, you will receive zero points for that period. When a choice is made, the chosen committee will automatically allocate 120 points among the four members. The outcomes may vary from one period to another, but are based on a theory of decision-making in committees. Once the allocation is made, you will immediately be shown the resulting allocation. At the end of the experiment, you will be paid according to your total earnings during the 60 periods, at an exchange rate of 1 point = 1 JPY.

partial-feedback treatment

There are a total of 60 periods. At each period, you have 30 seconds to make your choice between the two committees. If you do not make any choice within the 30 seconds at one period, you will receive zero points for the period. When a choice is made, the chosen committee will automatically allocate 120 points among the four members. The outcomes may vary from one period to another, but they are based on a theory of decision-making in committees. Once the allocation is made, you will be shown the number of points allocated to you. You will not see the allocations to the other members of the committee. At the end of the experiment, you will be paid according to your total points at an exchange rate of 1 point = 1 JPY.

no-feedback treatment

There are a total of 60 periods. At each period, you have 30 seconds to make your choice between the two committees. If you do not make any choice within the 30 seconds at one period, you will receive zero points for the period. When a choice is made, the chosen committee will automatically allocate 120 points between the four members. The outcomes may vary from one period to another, but they are based on a theory of decision-making in committees. You will not see the resulting allocation after each period. However, at the end of the experiment, you will be told the total points you have obtained during the 60 periods, and you will be paid according to the points earned over the 60 periods at an exchange rate of 1 point = 1 JPY.

If you have any questions, please raise your hand.

Appendix B: Subjects' Comments

In additional sessions, 10 subjects participated in sequence C $\rightarrow$ D. As shown in Figure 1, at least 6 subjects chose the correct option in Periods 41-45. Below are 6 answers to the post-experimental questionnaire of those subjects who succeeded in meaningful learning. The answers show that they had the proper reasoning for their choice, although the questionnaire was unfortunately not structured and the answers were written in a free format. Recall that in the additional sessions, however, the correct option is Choice 1 in the first 40 periods and it is Choice 2 in the subsequent 20 periods.

Question: Which option did you mainly choose? Why did you choose that option? Please explain the reason behind your choice.

- I realized that I could not obtain any reward without the approval of three voters including myself. Thus, in the second half, I chose the option in which it was less likely to be approved by two large voters only. (2 subjects)
- I chose the options that have more cases where three voters could win by themselves. (3 subjects)
- In Periods 1-40, there was a case where two large voters could collect 14 votes by themselves in Choice 2, but there was no such a case in Choice 1. Thus, I chose Choice 1. But, I sometimes chose Choice 2, because I was not sure about how 120 points would be distributed. In Periods 41-60, there were two cases where two voters exept me could collect 9 votes by themselves in Choice 1, but there was one case where two voter could collect 9 votes by themselves. Thus, I chose Choice 2 many times. But, I sometimes chose Choice 1, because I did not know how 120 points would be distributed. (1 subject)

References

Aleskerov, F., Beliani, A., Pogorelskiy, K. 2009. Power and preferences: an experimental approach. National Research University, Higher School of Economics, mimeo.

Arifovic, J., McKelvey, R.D., Pevnitskaya, S. 2006. An initial implementation of the Turing tournament to learning in two person games. Games Econ. Behav. 57, 93-122.

Banzhaf, J.F. 1965. Weighted voting doesn't work: a mathematical analysis. Rutgers Law Review 19, 317-343.

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

Cheung, Y.-W., Friedman, D. 1997. Individual learning in normal form games: some laboratory results. Games Econ. Behav. 19, 46-76.

Cooper, D.J., Kagel, J.H. 2003. Lessons learned: generalized learning across games. Amer. Econ. Rev. 93, 202-207.

——— 2008. Learning and transfer in signaling games. Econ. Theory 34, 415-439.

Deegan, J., and Packel, E. 1978. A new index of power for simple n-person games. Int. Jour. Game Theory 7, 113-123.

Dufwenberg, M., Sundaram, R., Butler, D. J. 2010. Epiphany in the game of 21. Jour. Econ. Behav. Org. 75, 132-143.

Erev, I., Roth A.E. 1998. Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. Amer. Econ. Rev. 88, 848-81

Erev, I., Ert, E., Roth, A.E. 2010. A choice prediction competition for market entry games: an introduction. Games 1, 117-136.

Esposito, G., Guerci, E., Lu, X., Hanaki, N., Watanabe, N. 2012. An experimental study on "meaningful learning" in weighted voting games. mimeo., Aix-Marseille University.

Felsenthal, D.S., Machover, M., 1998. The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes. Edward Elgar, London

Fischbacher, U. 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Exper. Econ. 10, 171-178.

Gelman, A., Katz, J. N., Bafumi, J. 2004. Standard voting power indexes do not work: an empirical analysis. British Jour Pol Sci, 34, 657 窶 674.

Gilboa, I., Schmeidler, D. 2001. *A Theory of Case-Based Decisions.* Cambridge University Press.

Grant, S., Meneghel, I., Tourky, R., 2017. Learning under unawareness. mimeo. https://ssrn.com/abstract=3113983

Guerci, E., Hanaki, N. Watanabe, N. 2017. Meaningful learning in weighted voting games: an experiment. Theory. Decis. 83, 131-153.

Guerci, E., Hanaki, N. Watanabe, N., Esposito, G., Lu, X. 2014. A methodological note on a weighted voting experiment. Soc. Choice Welfare 43, 827-850

Hu, Y., Kayaba, Y., Shum, M. 2013. Nonparametric learning rules from bandit experiments. Games Econ. Behav. 81, 215-231.

Meyer, R.J., Shi, Y. 1995. Sequential choice under ambiguity: intuitive solutions to the armed-bandit problem. Manag. Sci. 41, 817-834.

Montero, M., Sefton, M., Zhang, P. 2008. Enlargement and the balance of power: an experimental study. Soc. Choice Welfare 30, 69-87.

Nowak, M,A., Sigmund, K. 1993. A strategy of win stay, lose shift that outperforms tit-for-tat in the prisoner's-dilemma game. Nature 364, 56-58.

Ogawa, K., Osaki, Y., Kawamura, T., Takahashi, H., Taguchi, S. Fujii, Y., Watanabe, N. 2020. Conducting economic experiments at multiple sites: Subjects' cognitive ability and attribute information. Kansai University RISS Discussion Paper Series No. 86, June 2020.

Rick, S., Weber, R.A. 2010. Meaningful learning and transfer of learning in games played repeatedly without feedback. Games Econ. Behav. 68, 716-730.

Shapley, L.S., Shubik, M. 1954. A method for evaluating the distribution of power in a committee system. Amer. Polit. Sci. Rev. 48, 787-792.

Watanabe, N. 2014. Coalition formation in a weighted voting experiment. Japanese Jour. Electral Stud. 30, 56-67.