What Information Hinders or Promotes Subjects' Meaningful Learning in a Bandit Experiment for Weighted Voting?: Mouse Tracking Evidence

Naoki Watanabe*

*Keio Business School

December 16, 2024, Washington, D.C.
2024 IEEE Int'l Conference on Big Data

what we have done

subject experiment of learning

- Subjects were asked to choose one of two weighted voting games repeatedly, and their payoffs were determined by a latent stochastic payoff-generating function.
- The vote apportionments and quotas of those games were hidden from subjects in windows on their computer screens. We mouse-tracked what information they viewed on the screens.
- After subjects had experienced a binary choice problem in rounds for learning something, we examined whether those subjects increased the number of choosing the answer which would give a higher expected payoff in a similar but different binary choice problem in the subsequent rounds.

When people applied what they had learned in a situation to a similar but different one and their inference was correct, we would say that they learned something that underlies commonly in those situations meaningfully.

- **Result 1**: The feedback information on subjects' cumulative payoffs might promote meaningful learning, whereas the feedback information on their own current payoffs did not, or even hindered it.
- **Result 2**: It would be plausible that even if subjects paid more attention to their cumulative payoff, they would fail in meaningful learning when they chose the runs of options randomly, unlike algorithms in machine learning.

Remarks:

- Some recent papers reported that withholding payoff-related feedback information promoted subjects' meaningful learning of latent features on strategic interactions among them.
  - In other papers I have published, the same thing was observed in the situation examined in this paper, where there were no strategic interactions.
  - This paper found (1) what feedback information promotes or hinders subjects' meaningful learning; cumulative payoff and current payoff, respectively, and (2) search behaviors that hinder people's meaningful learning; random choice of runs, win-stay-lose-shift strategy
- People may think more deeply about something behind the situation they are faced with, when they are placed in a situation with few feedback information.

Suppose that you were a subject.

choice stage (30 sec.); no choice within 30 seconds → zero point.
Each number in red can be viewed by a mouse-click.

Round 1 out of 60 remaining time 16

Choose one of two committees. The quota is 14 in both committees.
The committee you chose allocates 120 points among four members.
Click on Choice 1 button or Choice 2 button within 30 seconds.

Choice 1	member	YOU	Player 2	Player 3	Player 4
	votes	V1	V2	V3	V4

Choice 2	member	YOU	Player 2	Player 3	Player 4
	votes	W1	W2	W3	W4

Figure: Choice stage.

feedback stage 1 (30 sec.); OK button → feedback stage 2. Each number in red can be viewed by a mouse-click.

Round 1 out of 60 remaining time 23

Click on OK button within 30 seconds after seeing the following information.
You chose the following committee.

Choice 1	member	YOU	Player 2	Player 3	Player 4
	votes	X1	X2	X3	X4

The committee decided to allocate 120 points in this time as follows.

member	YOU	Player 2	Player 3	Player 4
Points	p1	p2	p3	p4

OK

Figure: Feedback stage 1.

feedback stage 2 (30 sec.); OK button → next round. Each number in red can be viewed by a mouse-click.

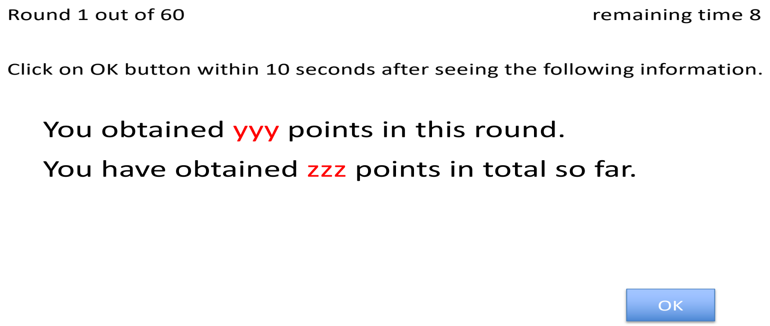


Figure: Feedback stage 2.

- Answer 60 binary choice problems (60 rounds).
  When you finish 60 rounds, you are asked to wait until the others finish.
- 1 point = 1 JPY (about 115 JPY/ USD in 2017), added to the payment for participation (1000 JPY)

mouse-tracking to capture what information subjects view on their monitor

- When measuring gaze with eye trackers, subjects often unconsciously move their eye-sight to places other than the measurement targets. (a kind of noise)

1. Introduction: Research Question

- **Question**: What information did subjects view, when they could (not) "meaningfully learn" the latent feature of weighted voting in a "2-armed bandit experiment"?
  - Felsenthal and Machover (1998): It would be difficult (for people) to see the underlying relationship between the actual voting powers and the nominal voting weights.
- When people generalize what they have learned in a situation to a similar but different one, this higher order concept of learning is called meaningful learning (Rick and Weber, 2010).
- 2-armed bandit experiment: Subjects choose one of two weighted voting games repeatedly. Payoffs are determined stochastically according to a payoff-generating function that is hidden from subjects.

1. Introduction: Previous Research

- Guerci et al. (2017, TD) could **not observe meaningful learning** by subjects who received immediate feedback information on their current payoffs.
  - **Withholding feedback information induced meaningful learning** (introspective thinking). Did the subjects get confused by the feedback information?
  - Ogawa et al. (2021, mimeo.) reconfirmed a similar result at four different experimental sites.
- ref. Guerci et al. (2014, SCW): subjects played 4-person weighted voting games
  - not focused on their learning but compared frequencies of their "mistakes" observed in bargaining protocols in voting. (single approval or multiple approval...)
  - Esposito et al. (2012, mimeo.): subjects chose a weighted voting game and then they played the game they chose. $\rightarrow$ subjects' learning was affected by others' voting behavior.

1. Introduction: Main Result

We measured how often or long each subject views information necessary for his or her choice without duplication of viewing time.

The feedback Information about cumulative payoffs promoted meaningful learning, whereas the information on current payoffs did not, or even hindered it.

- The latter part of this result may explain the reason why meaningful learning was not observed by Guerci et al. (2017), because in their experiment only the current payoffs were shown to the subjects.

search behavior: win-stay-lose-shift (WSLS), random choice of runs
$\rightarrow$ outcomes: current payoff and failure in meaningful learning

2. Experimental Design

- Subjects choose one of two weighted voting games.
  - $[q; v_1, v_2, v_3, v_4]$; q is the quota, v_i is the voting weight.
  - Each subject acts as Member 1, who has v_1.
  - clearly informed that the other three members are all fictitious.
- one binary choice problem (Problem) for the first 40 rounds, and a similar but different one in the following 20 rounds;
  - This information was not provided to subjects.
- four sequences: A $\rightarrow$ B, B $\rightarrow$ A, C $\rightarrow$ D, D $\rightarrow$C
- payoff-generating function: based on Deagan-Packel index
  e.g., Problem A, Choice 1 $\rightarrow$ 40 points with Prob=2/3,
  0 point with Prob=1/3

Table: Binary choice problems and expected payoffs for Member 1

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	$[9; \mathbf{1}, 3, 5, 6]$	$(120 \times 1/3 \times 1/3)$	$[9; \mathbf{1}, 2, 6, 6]$	$(120 \times 2/3 \times 1/3)$

choice stage (30 sec.); no choice within 30 seconds $\rightarrow$ zero point.
Each number in red can be viewed by a mouse-click.

Round 1 out of 60 remaining time 16

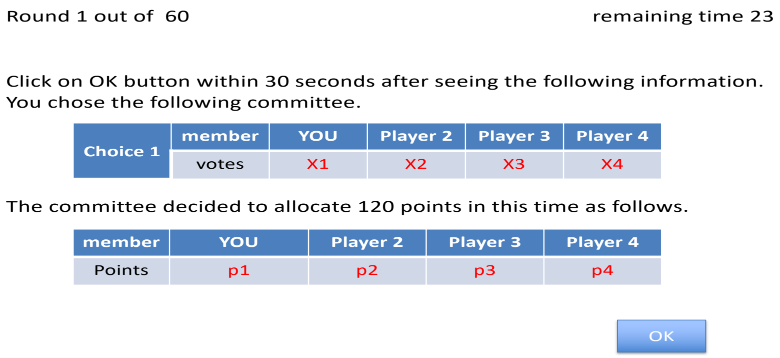
Choose one of two committees. The quota is 14 in both committees.
The committee you chose allocates 120 points among four members.
Click on Choice 1 button or Choice 2 button within 30 seconds.

Choice 1	member	YOU	Player 2	Player 3	Player 4
	votes	V1	V2	V3	V4

Choice 2	member	YOU	Player 2	Player 3	Player 4
	votes	W1	W2	W3	W4

Figure: Choice stage.

feedback stage 1 (30 sec.); OK button → feedback stage 2. Each number in red can be viewed by a mouse-click.



Round 1 out of 60 remaining time 23

Click on OK button within 30 seconds after seeing the following information.
You chose the following committee.

Choice 1	member	YOU	Player 2	Player 3	Player 4
	votes	X1	X2	X3	X4

The committee decided to allocate 120 points in this time as follows.

member	YOU	Player 2	Player 3	Player 4
Points	p1	p2	p3	p4

OK

Figure: Feedback stage 1.

feedback stage 2 (30 sec.); OK button $\rightarrow$ next round. Each number in red can be viewed by a mouse-click.

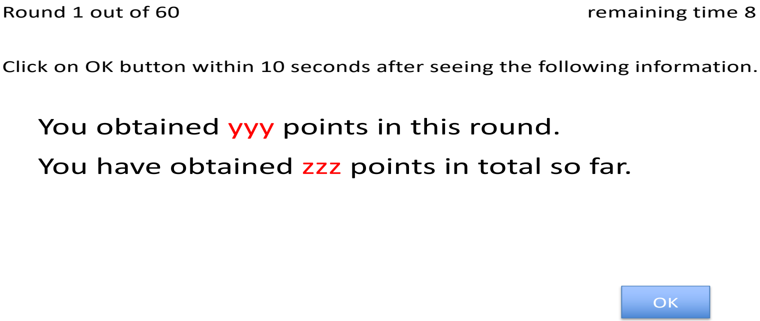


Figure: Feedback stage 2.

- Subjects answer 60 binary choice problems.
  $\rightarrow$ a message that asks them to wait until the others finish
  ($\rightarrow$ Questionnaire and Raven test)
- 1 point = 1 JPY, added to the payment for participation (1000 JPY)

3-1. Analysis: Session Details

- December 17 in 2014 to January 17 in 2017
- at the University of Tsukuba
- undergraduate students who did not know voting indices
- 40 subjects for each sequence (160 subjects in total)
- the average amount paid as a reward was 2507 JPY

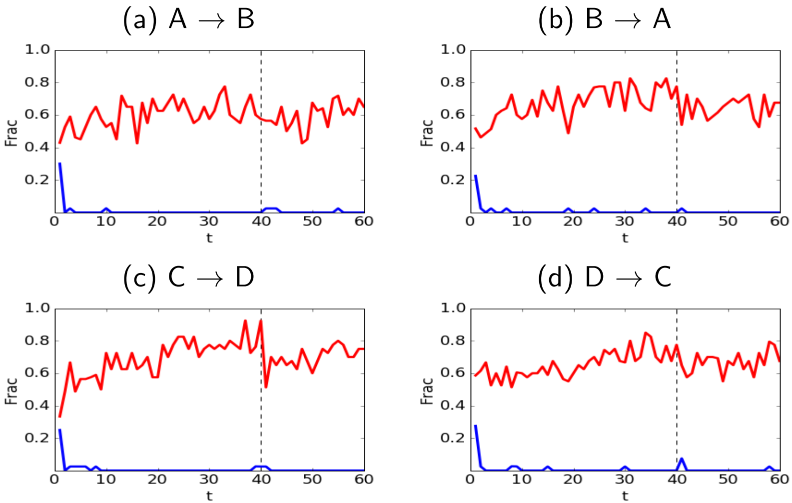


Figure: RED: fraction of subjects who chose the correct answer among those who made choices before the time limit. BLUE: fraction of those who failed to make their choice before the time limit.

- FR_k^i: the relative frequency of rounds in which subject i chose the correct answer within the k-th block of 5 consecutive rounds.
- The change in the relative frequencies that subject i chose the correct answer between the l-th block and the m-th block is defined as

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

  where $l > m$.
- $\Delta\mathrm{FR}_{l,m}$ $(\mathrm{FR}_l, \mathrm{FR}_m)$: a vector the i-th component of which is $\Delta\mathrm{FR}_{l,m}^i$ $(\mathrm{FR}_l^i, \mathrm{FR}_m^i)$.
  - When the elements of FR_l are, on average, significantly larger than those of FR_m, we write this as $\Delta\mathrm{FR}_{l,m} > 0$.

3-2. Results: Aggregate Data

Definitions: For each binary choice problem, we consider that subjects learned the "correct" answers if $\Delta\mathrm{FR}_{8,1} > 0$ was statistically confirmed and that those who learned in the binary choice problem meaningfully learned the "underlying structure" of weighted voting games if $\Delta\mathrm{FR}_{9,1} > 0$ in the same binary choice problem was statistically confirmed.

Result 1: For all binary choice problems, the subjects learned the correct answer.

Result 2: For all binary choice problems except Problem B, the subjects meaningfully learned the underlying structure of weighted voting.

	Problem A	Problem B	Problem C	Problem D
$\Delta\mathrm{FR}_{2,1}$	0.0415	**0.0086**	**0.0396**	0.6943
$\Delta\mathrm{FR}_{8,1}$	**0.0143**	$<$**0.0001**	$<$**0.0001**	$<$**0.0001**
$\Delta\mathrm{FR}_{9,1}$	**0.0147**	0.2498	$<$**0.0001**	**0.0230**

Table: P-values for the two-sided signed-rank test (normalized). The null hypotheses are $\Delta\mathrm{FR}_{2,1} = 0$, $\Delta\mathrm{FR}_{8,1} = 0$, and $\Delta\mathrm{FR}_{9,1} = 0$, respectively

- An auxiliary experiment was conducted at Kansai University: Cumulative payoffs were shown to the subjects' monitors as well as their current payoffs, and similar results to Results 1 and 2 were observed. $\rightarrow$ The use of mouse trackers would not affect subjects' choice in the aggregate level.

3-3. Results: Individual Data

- two groups of subjects $G_0^{91}(s)$ and $G_1^{91}(s)$ for binary choice problem $s = A, B, C, D$: $\Delta \mathrm{FR}^i_{9,1} = \mathrm{FR}^i_9 - \mathrm{FR}^i_1 > 0$ or not
- canonical discriminant analysis: the first 40 rounds to identify important mouse-tracked variables that had a large effect on the classification of the two groups.
  - Box's M test on variance-covariance matrices → Mahalanobis distance was applied to identify the discriminant function

variable	
A_count:	numbers of views of A; $A = v_r, w_r, x_r, p_r$ $(r = 1, 2, 3, 4)$, yyy, zzz
A_time:	cumulative time for viewing A
decision_time:	time spent for the final decision in the choice stage
A_decision_time:	relative length of time spent for viewing A up to the final decision
main2_ok:	time spent in feedback stage 1
main3_ok:	time spent in feedback stage 2
no_info:	dummy variable that takes a value of 1 if the individual subject views $v_1, \ldots, v_4$ or $w_1, \ldots, w_4$ even once; otherwise, 0
judgment:	dummy variable that takes a value of 1 if $\Delta \mathrm{FR}^i_{8,1} > 0$; otherwise, 0
judgment2:	dummy variable that takes a value of 1 if $\Delta \mathrm{FR}^i_{9,1} > 0$; otherwise, 0

Table: Major variables.

Definition: For each sequence of binary choice problems, we considered that subject i learned the "correct" answers if $\Delta\mathrm{FR}^i_{8,1} > 0$ and that those who learned in a binary choice problem meaningfully learned the "underlying structure" of weighted voting games if $\Delta\mathrm{FR}^i_{9,1} > 0$ in another binary choice problem in the sequence.

- the numbers of subjects who succeeded in learning (meaningful learning) were 20, 30, 32, and 27 (14, 13, 19, 17), respectively, under the following definitions.
- For every binary choice problem $s = A, B, C, D$, we could reject the null hypothesis of no difference between $G^{81}_0(s)$ and $G^{81}_1(s)$ (between $G^{91}_0(s)$ and $G^{91}_1(s)$) for independent variables, because Wilks' lambda < 0.001.

A	**zzz_count**	*p1_time*	*v4decision_time*	**yyy_count**	*p3_time*
	0.678	0.627	-0.486	-0.484	-0.378

B	*decision_time*	*x1_decision_time*	*x1_count*	**zzz_count**	**yyy_count**
	-0.827	-0.552	0.521	0.506	0.241

C	*x1_count*	*x3_count*	**yyy_time**	*w4_time*	*v3_time*
	-1.359	1.040	-0.549	0.543	0.473

D	*p4_time*	**zzz_count**	*p1_time*	*no_info*	**yyy_count**
	0.749	0.479	0.315	-0.298	-0.142

Table: Standardized coefficients in the canonical discriminant functions: the four largest absolute values observed in the first 40 rounds in the discriminant function that separated well $G_0^{91}(s)$ and $G_1^{91}(s)$ for binary choice problem $s = A, B, C, D$.

Result 3:

The information on subjects' cumulative payoffs promoted meaningful learning, whereas the information on their own current payoffs did not, or even hindered it.

- The latter part of this result may explain the reason why meaningful learning was not observed by Guerci et al. (2017), because in their experiment only the current payoffs were shown to the subjects.
  - Watanabe (2022): reconfirmation of meaningful learning without any feedback information (correct answers were reversed in the first 40 rounds and the following 20 rounds).

Subjects' Search Behavior 1

In what follows, we examine the null hypotheses in 20 rounds.

- We say that a subject engages in the win-stay-lose-shift (WSLS) strategy when he or she continues to choose the same answer immediately after obtaining 40 points, while he or she changes answers immediately after obtaining 0 point (Nowak and Sigmund, 1993).

For each individual subject i, let a_i denote the number of rounds in which $yyy\_count - zzz\_count \geq 0$ and let b_i denote the number of rounds in which $yyy\_time - zzz\_time \geq 0$.

- We say that in 20 consecutive rounds, subject i paid more attention to the current payoffs than to the cumulative payoffs, if either $a_i \geq 5$ or $b_i > a_i$ in those rounds.

Define *current payoff/WSLS* as the number of subjects who paid more attention to the current payoff among those who took the WSLS strategy and *meaningful/WSLS* as the number of subjects who succeeded in meaningful learning among those who took the WSLS strategy.

In the next table, the values in the parentheses are the numbers of subjects who did not pay more attention to the current payoff among those who took the WSLS strategy and the numbers who failed in meaningful learning among those who took the WSLS strategy.

	current payoff/WSLS			meaningful/WSLS		
rounds	1-20	21-40	41-60	1-20	21-40	41-60
A → B	4 (3)	5 (3)	**6 (1)**	4 (4)	5 (3)	4 (2)
B → A	5 (1)	3 (0)	2 (1)	2 (4)	2 (1)	2 (1)
C → D	**7 (0)**	4 (3)	1 (4)	2 (5)	2 (5)	1 (4)
D → C	**5 (0)**	**6 (0)**	**7 (0)**	**0 (5)**	**0 (6)**	2 (5)

Table: Frequencies of observations: current payoff, meaningful learning, and WSLS. The rejection of the null hypothesis at the 5% significance level in the one-sided binomial test is indicated in boldfaced value.

Result 4:
For Problem D in the sequence of binary choice problems D and C, subjects who took the WSLS strategy paid more attention to the current payoffs than to the cumulative payoffs and they failed in meaningful learning.

Subjects' Search Behavior 2

Define *current payoff/runs* as the number of subjects who paid more attention to the current payoff among those who chose the runs randomly and *meaningful/runs* as the number of subjects who succeeded in meaningful learning among those who chose the runs randomly.

In the next table, the values in the parentheses are the numbers of subjects who did not pay more attention to the current payoff among subjects who chose the runs of options randomly and the numbers of those who failed in meaningful learning among those who chose the runs randomly.

- Under the null hypothesis in the (Wald-Wolfowitz) runs test, the number of runs of options chosen by a subject is a random variable.
  - Computer scientists consider the optimal length of runs of options in which decision-makers should continue to choose the same options for accumulating the feedback information.

Result 5:

For Problem D in the sequence of binary choice problems D and C, subjects who chose the runs of options randomly paid more attention to the current payoffs than to the cumulative payoffs and they failed in meaningful learning.

	current payoff/runs			meaningful/runs		
rounds	1-20	21-40	41-60	1-20	21-40	41-60
A $\to$ B	**21 (7)**	13 (10)	15 (10)	9 (14)	8 (16)	**7 (16)**
B $\to$ A	**25 (4)**	17 (9)	**16 (6)**	13 (17)	14 (14)	**8 (15)**
C $\to$ D	**19 (1)**	**16 (4)**	**17 (5)**	10 (10)	9 (10)	12 (11)
D $\to$ C	**30 (0)**	**26 (2)**	**15 (2)**	11 (20)	**9 (20)**	**6 (18)**

Table: Frequencies of observations: current payoff, meaningful learning, and random choice of runs. The rejection of the null hypothesis at the 5% significance level in the one-sided binomial test is indicated in boldfaced value.

Rounds 21-40 are more important than rounds 1-20 due to the definition of learning.

- As far as Problem D was concerned, it was confirmed that the subjects who took the WSLS strategy paid more attention to the current payoffs and that those subjects failed in meaningful learning (Result 4).
- The same things were confirmed when the subjects chose the runs of options randomly (Result 5).

4. Final Remarks

Summary:

- Information on subjects' cumulative payoffs promoted subjects' deep understanding (meaningful learning) the underlying structure of weighted voting.
- The immediate feedback information on their own current payoffs did not promote their meaningful learning, or even hindered it.
  - Such information would prevent them from meaningfully learning under WSLS strategy or random choice of runs.
  - other search behavior in other binary choice problems?
- future research
  - length of memory $\rightarrow$ case-based decision theory (Gilboa and Schmeidler (1995, 2001)
  - response time in the process of learning (ref. Rubinstein, 2013)