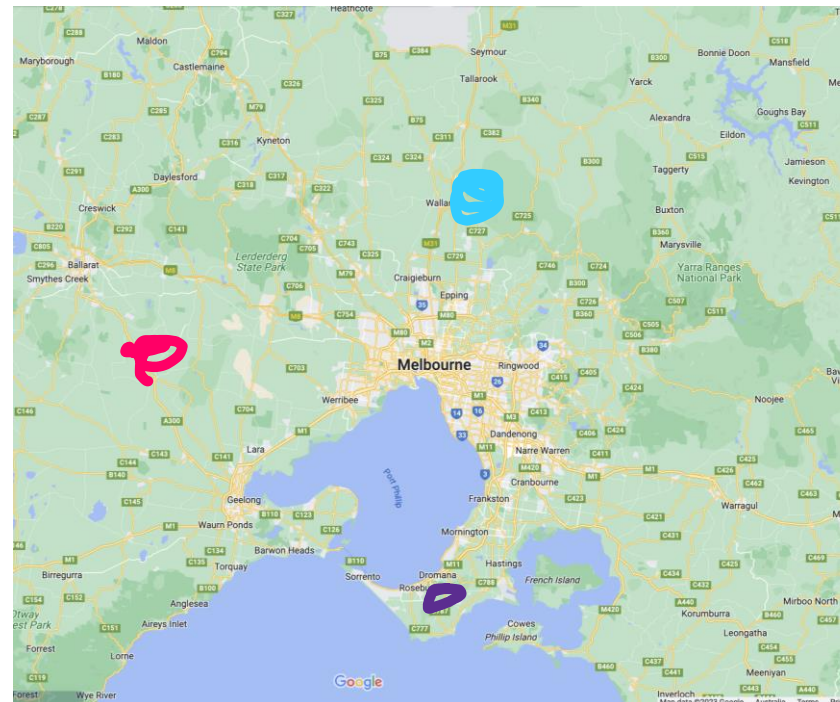


Human Collective Intelligence under Dual Exploration- Exploitation Dilemmas

Wataru Toyokawa, Hye-rin Kim, Tatsuya Kameda (2014)

Explore VS Exploit: Multi-Arm Bandit Task

- **Objective:** Meet wild white kangaroos (same or different) as many times as you can in 2023
- **Challenge:** Initially, the probability of finding a white kangaroo in each place is unknown
- **Constraint:** A limited # of Travels are allowed (100 travels per year)



Explore VS Exploit: Multi-Arm Bandit Task

- **Objective:** Maximize the sum of rewards
- **Challenge:** Initially, the reward probability of each option is unknown
- **Constraint:** Limited Trials (*And Computational Complexity...OF COURSE*)



Research Question

How can the availability of **social information** affect human collective **performance** on the MAB task in the face of within and between individual exploration-exploitation dilemmas?

Methods – 30-Armed Bandit Task

Current round: 2
Your total earnings: 130

		C6			

- Maximize cumulative performance within 100 trials
 - Direct Feedback (Reward) after choice (Always non-negative rewards)
 - Rewards are drawn from 6 category of uniform distributions
1. 11 of 30 bandit $X_1 \sim U(0, 15)$
 2. 8 of 30 bandit $X_2 \sim U(15, 30)$
 3. 5 of 30 bandit $X_3 \sim U(30, 45)$
 4. 3 of 30 bandit $X_4 \sim U(45, 60)$
 5. 2 of 30 bandit $X_5 \sim U(60, 75)$
 6. 1 of 30 bandit $X_6 \sim U(75, 90)$

Methods – Participants

- 194 Participants (42 Females) are divided into three Paradigms, each running 100 trials
- A and B have social information available, C does not

14 Groups of 5 Participants

A Frequency only

Current round: 2					
Your total earnings: 130					
			1		
	2			1	
		1			

14 Groups of 5 Participants

B Frequency-plus-evaluation

Current round: 2					
Your total earnings: 130					
			4.00 1		
	3.50 2			- 1	
		2.00 1			

54 Individual Participants

C No Collaboration

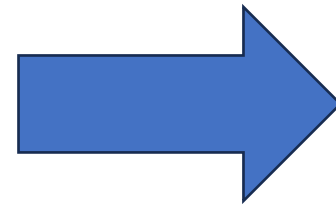
Current round: 2					
Your total earnings: 130					

Methods – Reformat Data

- Rewards are drawn from 6 category of

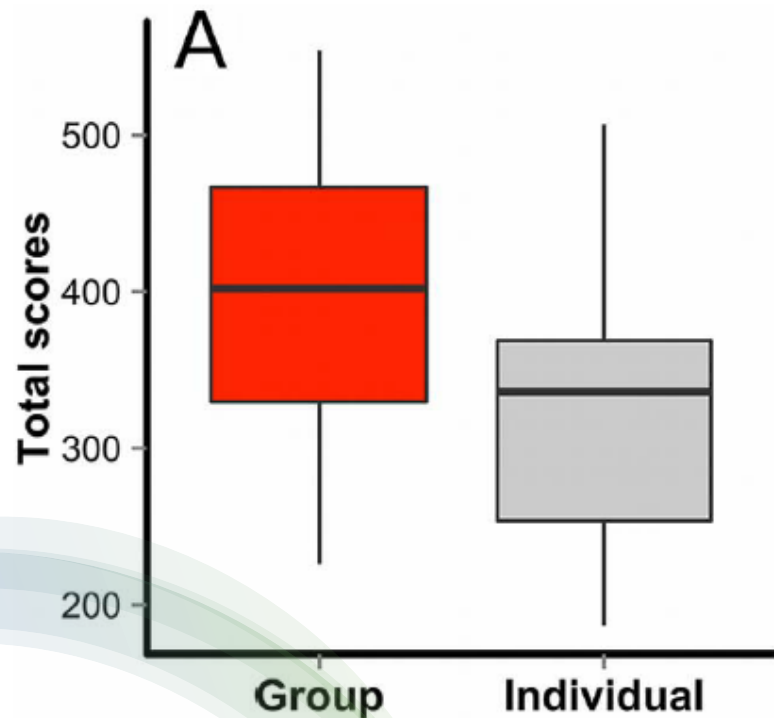
uniform distributions

- | | | | |
|----|-----------------|---------------------|---|
| 1. | 11 of 30 bandit | $X1 \sim U(0, 15)$ | 1 |
| 2. | 8 of 30 bandit | $X2 \sim U(15, 30)$ | 2 |
| 3. | 5 of 30 bandit | $X3 \sim U(30, 45)$ | 3 |
| 4. | 3 of 30 bandit | $X4 \sim U(45, 60)$ | 4 |
| 5. | 2 of 30 bandit | $X5 \sim U(60, 75)$ | 5 |
| 6. | 1 of 30 bandit | $X6 \sim U(75, 90)$ | 6 |



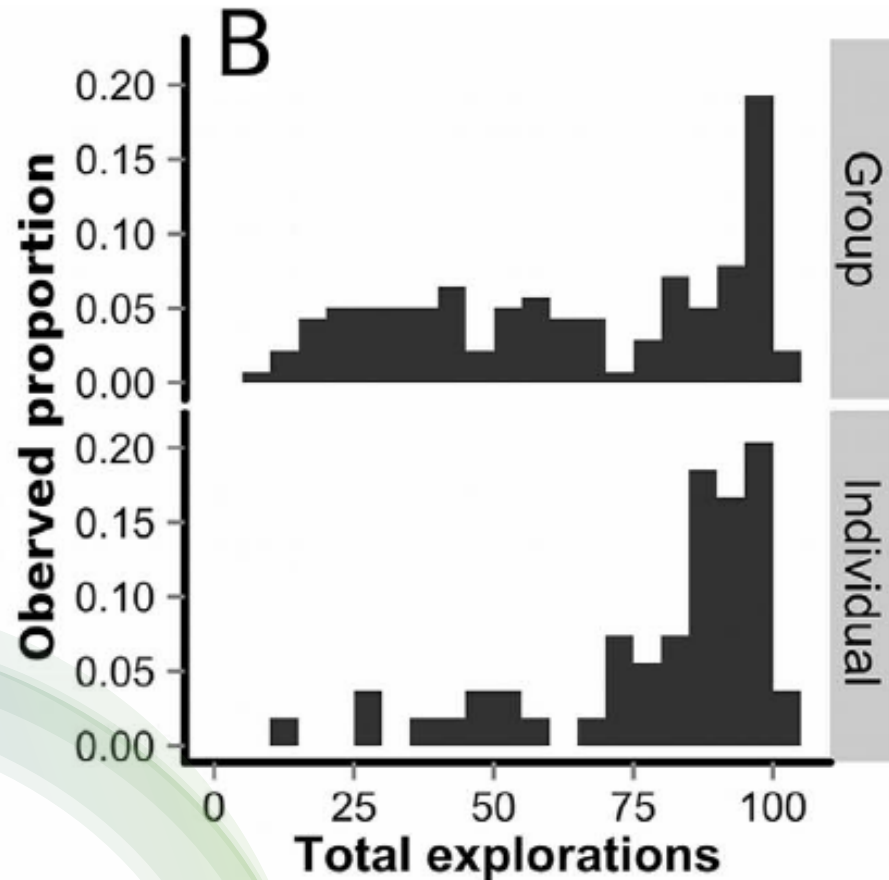
Randomness

Result – Social Information Vs Individual Box Plot



- Mean treatment group score is **397.3** and mean control group score is **321.7**. The difference is significant at 1% level
- Interestingly, the best performed individual achieved about 550 scores in the grouped condition!! That means, he/she must have selected the best patch of the 30 patches more than half of the 100 trails!

Result – Frequency of Exploring



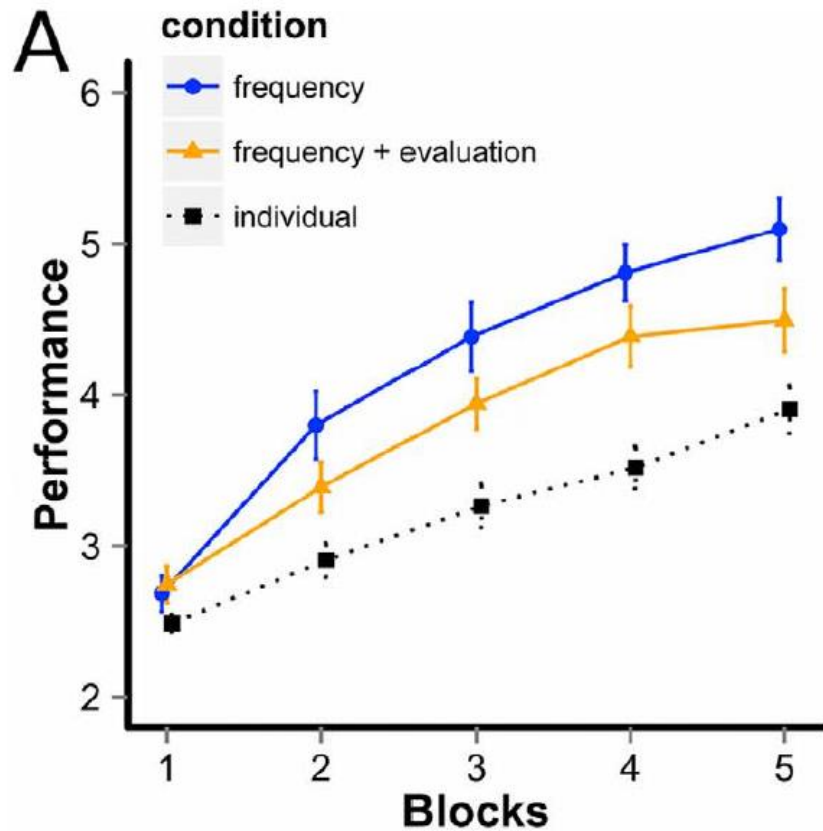
- Exploration is defined as a choice in which a participant selected options other than the “greedy” option that had yielded the largest cumulative payoff so far in her/his own experience
- **Participants engaged in fewer exploratory choices in the group condition** than in the individual condition (Significant – using Markov Chain Monte Carlo method)

Result – Mini summary 1

- Social information reduced each participant's total exploration time, helping her/him to exploit better options for a longer time in the group condition than in the individual condition

Exploration-Exploitation dilemma was more efficiently resolved in the group condition than in the individual condition

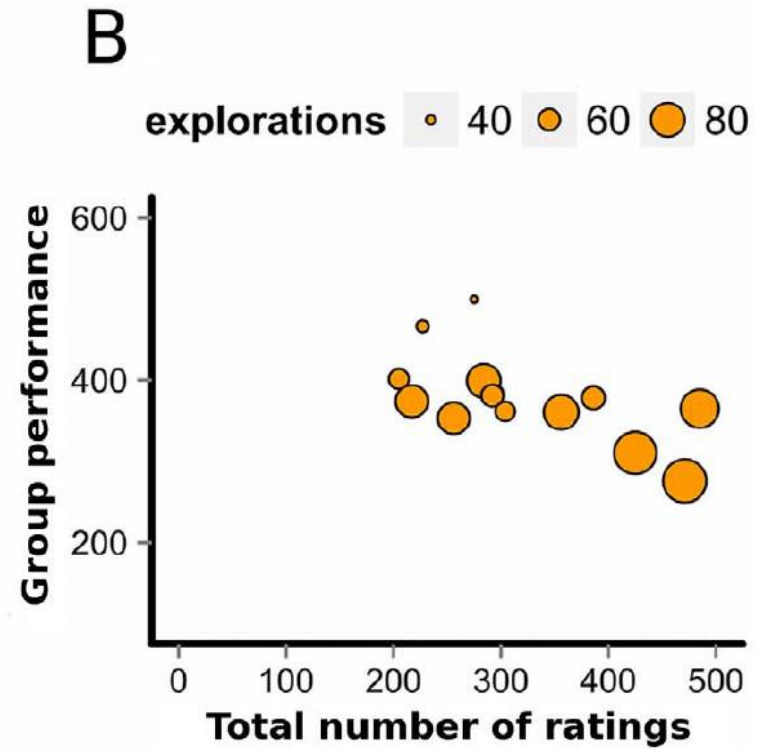
Result – Condition A Vs Condition B



- 20 trials as a block (N=14 groups for condition A, N=14 groups for condition B and N=54 for control group)
- Average performance increases over time (significant at 1%)
- The **overall average** performance of condition A is not significantly different from condition B (**insignificant** at 10%)
- The average **performance** of condition A is significantly **higher** than condition B in **block 5** (**significant** at 5%)

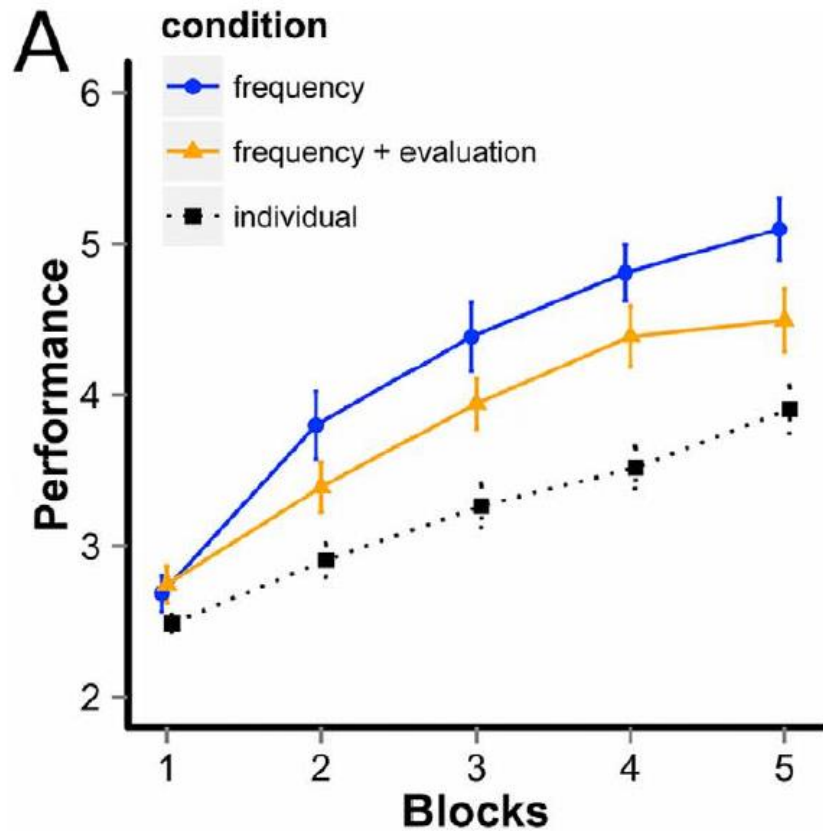
More info is detrimental to performance in later blocks?

Result – Condition B Rating and Performance



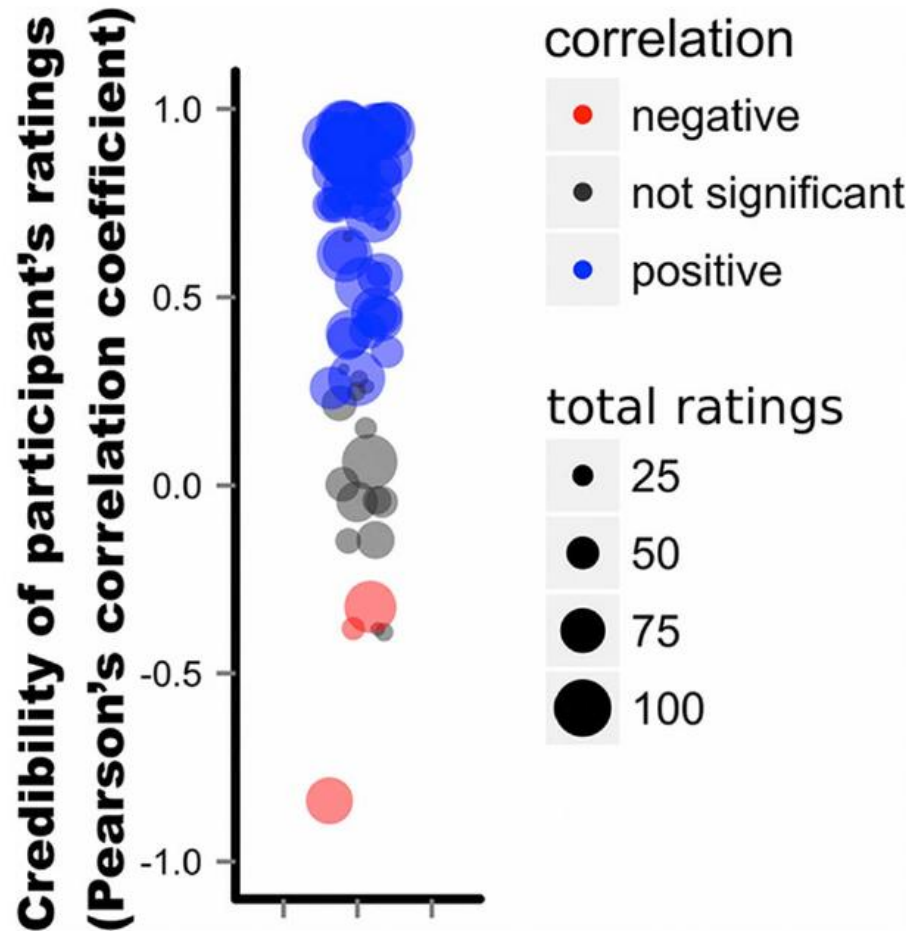
- Ratings are negatively correlated to group performance ($r = -0.613$, significant at 5%)
- Explorations are negatively correlated with group performance ($r = -0.854$, significant at 1%)

Discussion – Condition A Vs Condition B



- More information is always better? or **less is more?**
- **ISSUES:**
 - People may deliberately rate “good” patches “bad” to mislead others (although they have no monetary incentive to do so)
 - The rating is subjective

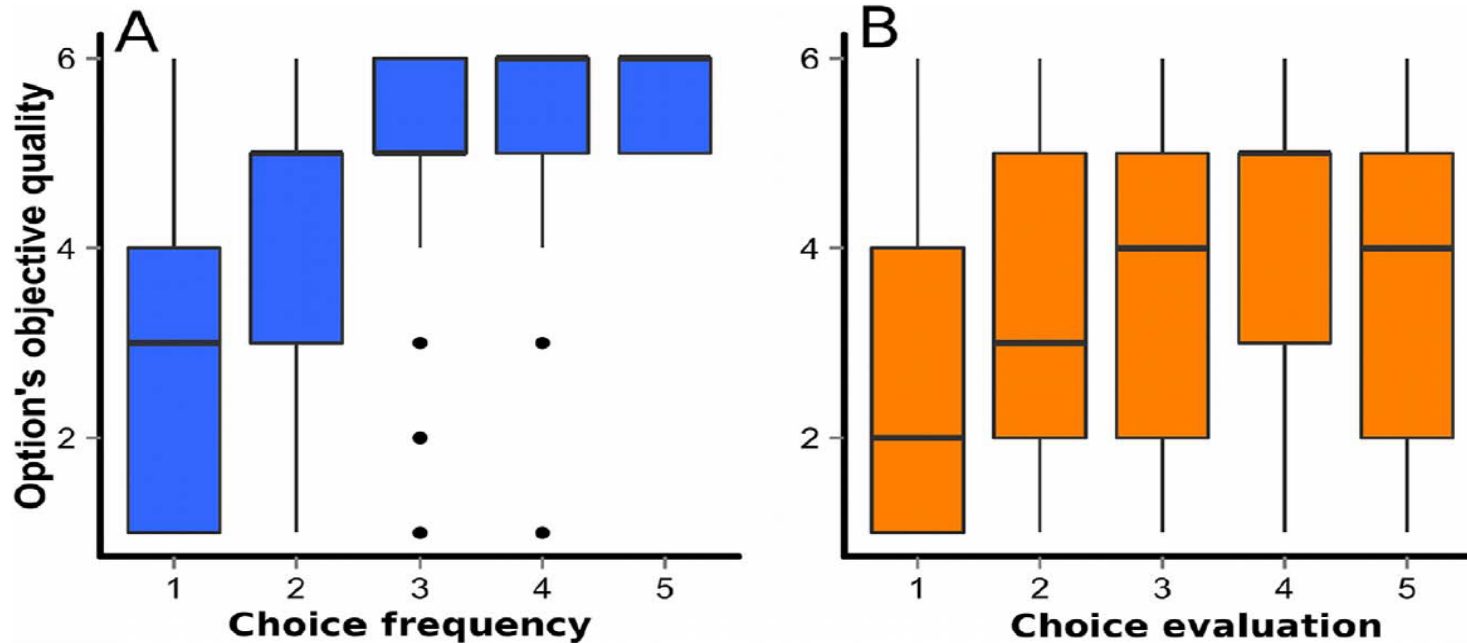
Discussion – Spiteful Ratings?



- 20 out of 70 participants provided invalid evaluations of options (Ratings not significantly positively correlated with their payoffs received)
- 3 people rate spitefully

Discussion – Validity of Social Info?

- Choice frequency is positively related to performance ($r = 0.45$, significant at 1%)
- Ratings is positively related to performance ($r = 0.22$, significant at 1%)
- **Choice frequency has a significantly higher correlation coefficient than ratings** (significant as 1%)



Result - Mini Summary 2

- People's choice history is a more valid source of information than people's ratings
 - When shopping, look at how many people bought this product in the past year, rather than look at the average ratings
- People perform better when they don't have access to the rating information
 - Maybe Amazon should consider this?
 - Facebook have adopted a "Like" button where only positive evaluations are allowed

The Facebook logo, consisting of the word "facebook" in a bold, blue, lowercase sans-serif font.The Amazon logo, consisting of the word "amazon" in a bold, black, lowercase sans-serif font, with a curved orange arrow underneath it pointing from the 'a' to the 'z'.