Title
Comparing Human Predictions from Expert Advice to On-line Optimization Algorithms

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Xia, Feng
Zhu, Jianqiao
Griffiths, Tom

Publication Date
2023

Peer reviewed

Comparing Human Predictions from Expert Advice to On-line Optimization Algorithms

Feng Xia (feng.xia@princeton.edu)
Department of Computer Science
Princeton University

Jian-Qiao Zhu (jz5204@princeton.edu)
Department of Computer Science
Princeton University

Thomas L. Griffiths (tomg@princeton.edu)
Departments of Psychology and Computer Science
Princeton University

Abstract

On-line decision problems – in which a decision is made based on a sequence of past events without knowledge of the future – have been extensively studied in theoretical computer science. A famous example is the Prediction from Expert Advice problem, in which an agent has to make a decision informed by the predictions of a set of experts. An optimal solution to this problem is the Multiplicative Weights Update Method (MWUM). In this paper, we investigate how humans behave in a Prediction from Expert Advice task. We compare MWUM and several other algorithms proposed in the computer science literature against human behavior. We find that MWUM provides the best fit to people's choices.

Keywords: on-line problems; decision-making; resource-rationality; cognition

Many problems we face in our daily lives require us to make decisions about the future based on only our knowledge of the past. We may only discover our decision is suboptimal in years to come. Since we cannot return to the past and remake the decisions, how should we make decisions in the present so that our regret in the future will be minimized? This is the question explored in the computer sciencce literature studying on-line optimization problems (as opposed to off-line problems, in which we already have all the information) (e.g., Helmbold, Schapire, Singer, & Warmuth, 1998) and in the psychological literature on human decision-making (e.g., Hogarth, 1977; Rieskamp & Otto, 2006).

A famous example of an on-line optimization problem in computer science is Prediction from Expert Advice (Vovk, 1995; Cesa-Bianchi & Lugosi, 2006). For a period of T trials, one is asked to choose from m options based on advice from n experts. No other information is supplied except for the experts' advice. At the end of each trial, an outcome is revealed and each option incurs a cost. The goal is to minimize the overall decision costs over T trials. One simple and effective solution to the problem is the Multiplicative Weights Update Method (MWUM) (Arora, Hazan, & Kale, 2012). MWUM has been extensively used and studied in machine learning, game theory, and linear programming (Freund & Schapire, 1997; Blum, 2005; Foster & Vohra, 1999; Clarkson, 1988). Despite an intuitive resemblance between the algorithm and human decision-making, MWUM has not been investigated as an account of human behavior on this task.

To understand how humans behave in this on-line optimazation setting and compare their performance to that of algorithms that have been proposed in computer science, we designed experiments simulating the Prediction from Expert Advice problem. Participants made a series of investment decisions based on suggestions provided by a set of financial advisors. The decision problems were constructed so as to be diagnostic of the strategies that people use to solve this problem, allowing us to compare their performance with several algorithms from the optimization literature.

Our results showed that MWUM provided the most accurate account of the human data of the six models we examined. Given that MWUM is one of the optimal solutions to the problem (Arora et al., 2012), this result is consistent with previous work using optimization-based frameworks to account for human cognition (e.g., Anderson, 1990; Marr, 2010). These results provide a connection between theoretical computer science and cognitive science that may be a productive source of new models of human cognition.

On-line Decision-Making

In an on-line decision-making task, an agent is required to generate an immediate response without knowledge about the future. The response will then incur a cost contingent upon a future outcome. For example, in the Prediction from Expert Advice task a prediction is made based on the guidance of n experts. Practical examples of similar tasks include predicting whether it will rain today based on information from different forecasts and identifying opportunities for investment based on input from different financial advisors. The overarching objective is to minimize errors in prediction, approximating the performance of the most proficient expert.

Similar problems have been explored in the psychological literature on reinforcement learning. Since the early days of behaviorism, psychologists have explained human behavior via mechanisms that adjust our decisions according to rewards and punishment from our environment (Skinner, 1966). Feedback is only delivered after a behavior is chosen. From this perspective, the opinions of the experts could be treated as cues that become associated with outcomes. This process was formalized in the Rescorla-Wagner model, demonstrating how learning can occur via the adjustment of weights assigned to cues (Rescorla, 1972).

The Prediction from Expert Advice problem also shares some similarities with the multi-armed bandit problem that has been a focus of more recent work on reinforcement learning (e.g., Sutton & Barto, 2018). The key difference lies in

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

the fact that the rewards obtained by all the experts are visible to the player after each round, while in the multi-armed bandit problem the player can only observe the reward of the arm that they have chosen to play. This means there is no explore-exploit problem, which is a critical part of the multi-armed bandit problem (e.g., Thompson, 1933; Lai, Robbins, et al., 1985; Auer, Cesa-Bianchi, & Fischer, 2002).

Many of the algorithms that computer scientists have proposed to solve the Prediction from Expert Advice problem take the approach of allocating a distinct weight to each expert, denoted by $w_1, w_2, ..., w_n$, and forming predictions based on a weighted average of the experts' predictions. At the end of each trial, when the outcome is revealed, the algorithm modifies the weight w_i of expert i, decreasing it if a mistake is made or increasing it if the prediction is correct. As a direct comparison of these algorithms with human decision-making has not been explicitly conducted, our objective is to devise an experiment that facilitates the examination of human behavior in relation to this task, as well as its comparison to the various algorithms. In doing so, our goal is connecting the theoretical computer science literature with human behavior rather than exhaustively testing different psychological accounts of how people might be performing this task. We thus confine our analysis to algorithms related to those proposed in computer science, rather than considering a broader range of models, but we do highlight connections to the psychological literature when presenting these algorithms.

The Problem

The version of the Prediction from Expert Advice problem we analyze can be formalized as follows. For a trial t ranging from 1 to T, a binary decision must be made (e.g., choosing between "up" and "down"), contingent upon a collection of n expert opinions that provide predictions regarding the binary event. After choosing to adhere to an expert's prediction, the actual outcome is unveiled, and each expert's prediction incurs a cost, denoted as $m_i^{(t)}, \forall i \in \{1, 2, \ldots, n\}$. If the expert is incorrect this cost is 1, and if they are correct it is 0. The primary objective is to minimize the cumulative cost associated with decisions made across a total of T trials.

Considering the absence of performance guarantees for the experts (all experts may exhibit bad performance), optimality in this scenario is often characterized by attaining a performance level comparable to that of the most accurate expert instead of aiming for zero mistakes. Keeping this optimality criterion in mind, we now proceed to examine six potential algorithms for the Prediction from Expert Advice problem.

Algorithms

The Simple Majority Method (SSM) This algorithm simply follows the majority prediction at the current trial without using any historical information. For example, should 3 out of 5 experts predict "up" (constituting the majority), while the remaining 2 predict "down" in the current trial, then the method would yield a prediction of "up".

The Randomized Simple Majority Method (RSMM) RSMM represents a stochastic variant of SMM. Rather than deterministically adhering to the majority prediction, it selects a prediction with a probability proportional to the number of experts advising that prediction at the current trial. Once more, RSMM does not incorporate the experts' past performance. Using the same example used in SMM, the method would predict "up" with a probability of $\frac{3}{5}$ and predicts "down" with a probability of $\frac{2}{5}$.

The Weighted Majority Method (WMM) First proposed by Littlestone and Warmuth (1994), the WMM maintains a weighting of experts and returns the weighted majority prediction in each trial. Upon mistakes and successes, the method decreases and increases experts' weights, respectively, via a multiplicative update. The full algorithm is:

- Initialization: Fix a step-size parameter $\eta \leq \frac{1}{2}$. For each expert i, associate the weight $w_i^{(1)} = 1$
- For $t = 1, 2, ..., T$

   1. Predict "up" or "down" depending on which prediction has a higher total weight of experts advising it:

$$(\text{Predict "up" if } \sum_{\text{up-pred experts } i} w_i^{(t)} \geq \sum_{\text{down-pred experts } j} w_j^{(t)})$$

   2. Observe the costs of the experts $\mathbf{m}^{(t)}$
   3. For every expert i, update their weight:

$$w_i^{(t+1)} = w_i^{(t)}(1 - \eta m_i^{(t)})$$

The update rule in Step 3 says if an expert makes an error (i.e., $m_i^{(t)} = 1$), their weight is reduced.

The Multiplicative Weights Update Method (MWUM) MWUM is a randomized version of WMM. Instead of deterministically making a weighted majority prediction, it makes a prediction with probability proportional to the weights of the experts advising each option. It resolves the worst-case scenario in WMM when one option receives slightly higher weight than the other. The full algorithm is as follows:

- Initialization: Fix a step-size $\eta \leq \frac{1}{2}$. For each expert i, associate the weight $w_i^{(1)} = 1$
- For $t = 1, 2, ..., T$

   1. Predict "up" with probability p_{up} and predict "down" with probability p_{down}, where

$$p_{up} = \frac{\sum_{\text{up-pred experts } i} w_i^{(t)}}{\sum_{\text{up-pred experts } i} w_i^{(t)} + \sum_{\text{down-pred experts } j} w_j^{(t)}}$$

$$p_{down} = 1 - p_{up}$$

   2. Observe the costs of the experts $\mathbf{m}^{(t)}$
   3. For every expert i, update their weight:

$$w_i^{(t+1)} = w_i^{(t)}(1 - \eta m_i^{(t)})$$

It has been shown that the randomized version, MWUM, performs better than MWW by a factor of 2 (Arora et al., 2012). This procedure also has a meaningful connection to Bayesian inference and models based upon it such as Thompson sampling, as discussed in the Appendix.

The Deterministic Correct-rate Method (DCM) While keeping a weighting of experts, we now introduce an alternative update rule. Rather than being multiplicative, we propose the Deterministic Correct-rate Method (DCM), in which experts' weights are equal to the rate at which they have been correct so far. This can be seen as an additive update rule, similar to the Rescorla-Wagner rule (Rescorla, 1972), since we are updating the number of correct decisions for each expert in the numerator while the denominator remains the same for all experts. The full algorithm is as follows:

- Initialization: Associate the weight $w_i^{(1)} = 1$ for all experts
- For $t = 1, 2, \dots, T$
  1. Predict "up" or "down" depending on which prediction has a higher total weight of experts advising it:

$$(\text{Predict "up" if } \sum_{\text{up-pred experts } i} w_i^{(t)} \geq \sum_{\text{down-pred experts } j} w_j^{(t)})$$

  2. Observe the actual outcome. For every expert i, update their weight:

$$w_i^{(t+1)} = \frac{\#\text{ correct predictions up until } t}{t}$$

The Randomized Correct-rate Method (RCM) Motivated by preceding stochastic adaptations, we establish RCM in a similar way. The full algorithm is as follows:

- Initialization: Associate the weight $w_i^{(1)} = 1$ for all experts
- For $t = 1, 2, \dots, T$
  1. Predict "up" with probability p_{up} and predict "down" with probability p_{down}, where

$$p_{up} = \frac{\sum_{\text{up-pred experts } i} w_i^{(t)}}{\sum_{\text{up-pred experts } i} w_i^{(t)} + \sum_{\text{down-pred experts } j} w_j^{(t)}}$$

$$p_{down} = 1 - p_{up}$$

  2. Observe the true result. For every expert i, update its weight:

$$w_i^{(t+1)} = \frac{\#\text{ correct predictions up until } t}{t}$$

Testing the Algorithms

In order to compare these algorithms with human behavior, we devised a behavioral experiment based on the Prediction from Expert Advice Problem. To render the problem more tangible, we concentrated on the practical issue of stock prediction. Participants engaged in a sequence of decisions across a 10-day period, exclusively possessing access to the consultations of five experts: Alice, Bob, Chris, Diana, and Evan (we incorporated both conventional male and female names to circumvent potential gender stereotypes). Endowed with 10 points, participants would gain or lose one point for each correct or incorrect choice, respectively. Each day, a table recording the experts' performance in preceding trials was also displayed, permitting participants to consider the experts' historical performance if desired. This design aimed to help relax participant's memory constraints. To more effectively differentiate between the proposed models, we constructed multiple conditions targeting distinct model comparisons. A screenshot of the experiment is shown in Figure 1.

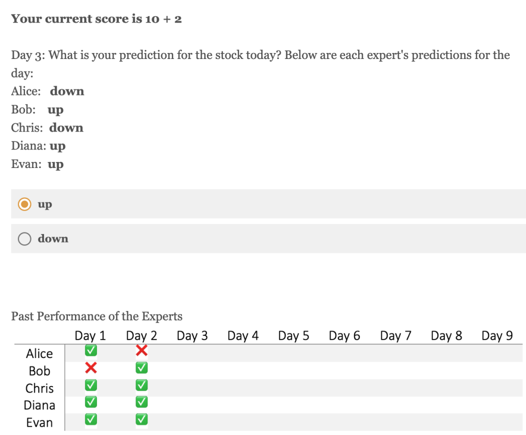

Figure 1: Screenshot of the experiment. Experts' advice for Day 3 was displayed on the top along with the participant's cumulative score. In addition, the historical record (Day 1 and 2) of each expert was shown in the bottom panel.

Methods

Participants A total of 120 participants (30 assigned to each of the four conditions) were recruited. These participants were chosen at random from a pool of Prolific users who resided in the United States, exhibited an approval rate exceeding 95%, and had completed over 100 submissions. Participants were compensated at an hourly rate of $12, and additional bonuses were awarded contingent upon their final score at the end of the experiment.

Stimuli Our goal was to develop sequences of expert advice and actual outcomes to maximally differentiate between the predictions generated by the algorithms. In order to gauge the capacity of a sequence to distinguish between models, we introduced a qualitative measurement. A *diagnostic difference* (dd) between the two models arises when one model predicts "up" with a probability greater than $50 + \alpha\%$, while the other predicts "up" with a probability less than $50 - \alpha\%$, where α is a variable that modulates the extent of the gap.

Table 1: Sequences of expert advice and true outcomes used in the experiment.

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	Correct rate
Alice	down	up	down	up	up	down	up	up	up	up	50%
Bob	down	up	up	up	up	up	up	up	up	up	50%
Chris	down	up	up	down	down	up	down	down	up	up	50%
Diana	up	down	down	down	up	down	down	up	down	down	50%
Evan	down	up	up	up	down	up	up	up	down	up	50%
True (Condition 1)	**up**	**down**	**down**	**down**	**down**	**up**	**up**	**up**	**up**	**up**	
Alice	down	up	down	down	down	up	down	down	up	down	50%
Bob	down	down	down	down	down	down	up	down	down	down	50%
Chris	down	down	down	down	down	down	up	down	down	down	50%
Diana	down	down	down	down	down	down	up	down	down	down	50%
Evan	up	up	up	up	up	up	down	up	up	up	50%
True (Condition 2)	**up**	**up**	**up**	**up**	**down**	**down**	**down**	**down**	**down**	**down**	
Alice	down	down	down	down	down	up	down	down	down	up	50%
Bob	up	up	up	up	down	down	down	down	down	up	60%
Chris	down	up	down	down	down	down	down	down	down	up	50%
Diana	down	up	up	up	down	down	down	down	down	up	70%
Evan	down	up	up	up	up	up	up	up	up	down	70%
True (Condition 3)	**down**	**up**	**up**	**up**	**up**	**up**	**up**	**down**	**down**	**up**	
Alice	down	up	up	down	down	down	down	up	up	up	60%
Bob	up	up	down	down	up	down	down	up	up	down	60%
Chris	up	up	up	down	down	down	down	up	down	up	60%
Diana	up	up	up	up	up	up	up	down	down	down	60%
Evan	down	down	up	up	down	down	down	up	up	up	60%
True (Condition 4)	**up**	**up**	**up**	**up**	**up**	**up**	**down**	**up**	**up**	**up**	

The ideal sequence would maximize the number of days on which each pair of models have a diagnostic difference. Note that MWUM, RSMM, and RCM are the randomized versions of WMM, SMM, and DCM, respectively. Hence, if a setup can maximally differentiate among these algorithms it can effectively distinguish among their deterministic counterparts, and hence all six models. Not incorporating historical information in the algorithm, RSMM is easily separable from MWUM and RCM. However, RCM and MWUM are intrinsically similar in their model predictions because both models perform similar weight updates. We therefore focused on finding sequences that differentiate MWUM and RCM.

Specifically, we randomly simulated 10^8 sequences with the gap parameter $\alpha = 10$. We set the step-size parameter $\eta \approx \sqrt{\frac{\ln n}{T}} = \sqrt{\frac{ln5}{10}} \approx 0.4$ because this value provides provably the tightest upper bound of regret for MWUM (Arora et al., 2012). One iteration of the procedure for finding the best sequence is as follows:

1. Randomly generate a 10-day ground truth sequence: "up" with probability $\frac{1}{2}$ and "down" with probability $\frac{1}{2}$.[1]

2. Randomly assign a correct rate to each of the five experts. [2] The correct rates were uniformly sampled from the following set $\{50\%, 60\%, 70\%, 80\%, 90\%\}$. [3]

3. For each expert, randomly select a proportion of days that also matches the correct rate for the expert and copy the true outcomes of those days as the expert's advice. Fill the remainder of the days with the opposite to ground truth.

Using this approach, we obtained four sequences that have the most diagnostically different days among MWUM, RSMM, and RCM: conditions 1 (dd = 5) and 2 (dd = 5) contrast RCM and RSMM, and conditions 3 (dd = 5) and 4 (dd = 4) contrast MWUM and RCM (see Table 1).

Procedure In each condition, for a 10-day period, participants were asked to make a prediction of whether the stock would go up or down given predictions suggested by the five experts (Alice, Bob, Chris, Diana, Evan). On each day, a table recording the past performance of the experts was provided. We used different colors for correct (green tick) and incorrect (red cross) predictions as visual aids supplied to the participants. Participants won one point for each correct decision and lost one point for each incorrect decision. Data were automatically collected by Qualtrics.

Results

We compare group-level data against the six on-line decision-making algorithms, assessing the percentage of participants who choose "up" (see Figure 2). To simplify the analysis, we assume that decisions made on each day are independent. As a result, our analysis considers each day's data as distinct data points in model comparison. To quantify the correspondence between model predictions and human data, we employ two metrics: Bayesian Information Criterion (BIC) and Pearson correlation coefficient (Pearson's r). Note that only WMM

[1] To reduce the impact of prior beliefs on stock behavior and rely solely on expert advice, the actual movement of the stock is not predictable from its past movements.

[2] An odd number of experts was chosen instead of an even number to prevent "tie" scenarios.

[3] Correct rates are always $\geq 50\%$ to prevent scenarios where poor experts (i.e., those with $< 50\%$ correct rates) are considered as valuable as good experts, as participants could potentially switch their recommendations to obtain similar payoffs.

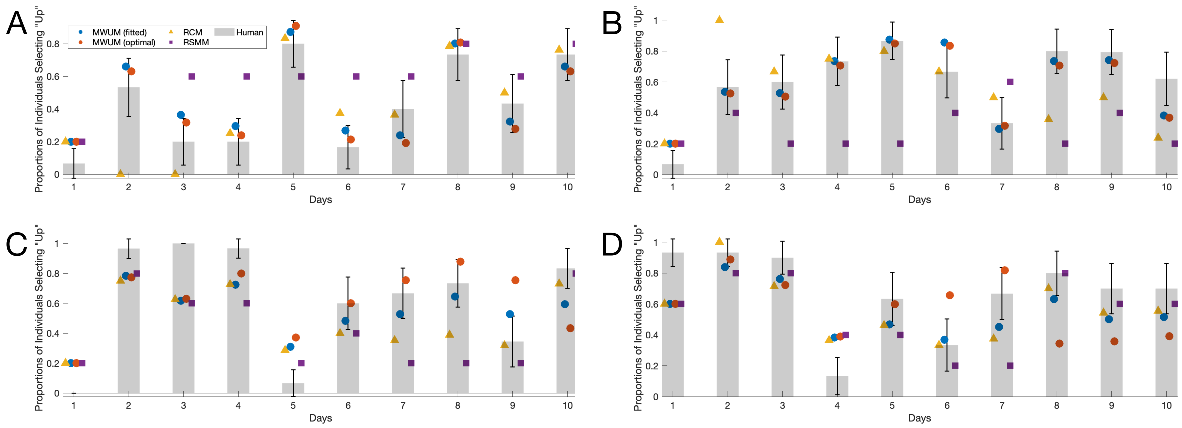

Figure 2: Comparing the proportions of individuals choosing "up" (represented by grey bars) with stochastic models including the Multiplicative Weights Update Method (MWUM) in blue (best-fitting η) and red (optimal $\eta \approx 0.4$) circles, the Randomized Correct-rate Method (RCM) in yellow triangles, and the Randomized Simple Majority Method (RSMM) in purple squares. Conditions 1 to 4 of our experiment appear in panels (A) to (D). Errorbars denote 95% confidence intervals across participants.

and MWUM require a choice of the step-size parameter η while the other four models have no free parameter. For η in WMM, we set $\eta \approx 0.4$, the value that minimizes the upper bound of the number of mistakes (Arora et al., 2012). For η in MWUM, we consider both the optimal $\eta \approx 0.4$ and the η that maximized the likelihood of human data.[4]

In aggregate, participants' behavior systematically deviated from complete randomness. Moreover, stochastic models tend to better correspond with human behavior (see Table 2). We thus focus on the comparison among stochastic models (i.e., MWUM, RCM, and RSMM). Figure 2 shows that MWUM and RCM are more indicative of empirical data than RSMM, suggesting people took into account the past performance of the experts. In addition, MWUM with $\eta = 0.3454$ in condition 1, 0.4254 in condition 2, 0.2384 in condition 3, 0.1707 in condition 4 fit via maximum likelihood best explains human data. Out of the four conditions, the best-fitting MWUM only has 1 diagnostically different day when comparing its predictions with human data, whereas RCM has 5 and RSMM has 17 diagnostically different days.

Given the increased complexity and the presence of an additional free parameter in MWUM compared to the other algorithms examined, we performed a quantitative model comparison using the Bayesian Information Criterion (BIC). The BIC balances accuracy and complexity by penalizing more intricate models. For simplicity, the BIC calculations were excluded for deterministic models (i.e., SMM, WMM, and DCM) as they need additional assumptions for likelihood calculations with discrete data. Table 2 reveals that the MWUM displayed the lowest BIC values, thereby designating it as the most preferable model. A similar conclusion can be drawn when examining the Pearson's r of the six models. It is noteworthy that the fitted MWUM offers a robust representation of human data, with Pearson's $r > 0.82$ ($p < .01$) across all four conditions, indicating a strong correlation.

Discussion

We face many on-line problems in our daily lives. For example, we have to decide whether to turn left or right at an intersection without knowing which route has less traffic and takes us home faster. So how do we make reasonably good decisions only knowing about the past but not the future? On-line optimization has been a popular area of research in theoretical computer science, resulting in algorithms such as the Multiplicative Weights Update Method (MWUM). In this paper, we compared the predictions of algorithms from the computer science literature against human behavior on the Prediction from Expert Advice problem, a classic problem in this literature. Our findings indicate that MWUM outperforms the other five algorithms we analyzed in terms of accounting for human behavior.

The analysis we presented here has several limitations. We acknowledge that our definition of "correct-rate" used in DCM and RCM is fairly arbitrary. One might, for example, argue that including a smoothing factor to the "correct-rate" will affect the results. In future work, we wish to explore more variations of models based on correct-rates, or more generally, additive rather than multiplicative rules for weight updates. We also recognize that the performance table we provided in the study could potentially bias the results, as it may encourage individuals to favor MWUM over other algorithms. In future studies, it may be worthwhile to explore different ways of presenting the past performance of experts in a manner that makes MWUM less salient. In particular, replicating the experiment under different conditions can help to validate our findings and verify if the performance table has a decisive impact on the strategies adopted by the participants.

[4]This choice of η is not optimal in the sense that it always renders the smallest error; it is optimal because it achieves least upper bound on the regret. In practice, the best η depends on the actual sequence.

Table 2: Summary of model fitting results.

	Condition 1		Condition 2		Condition 3		Condition 4	
Algorithm	BIC	Pearson's r	BIC	Pearson's r	BIC	Pearson's r	BIC	Pearson's r
MWUM (fitted)	**351.44**	**0.91*****	**354.34**	**0.88*****	**306.28**	**0.92*****	**352.58**	**0.82*****
MWUM (optimal)	-	0.90***	-	0.87***	-	0.65**	-	0.32**
RCM	4003.72	0.75**	4215.40	0.41**	4732.20	0.87***	3887.19	0.76**
RSMM	4996.94	0.69***	5543.76	-0.21*	5361.35	0.74**	4703.20	0.69**
SMM	-	0.47*	-	-0.39*	-	0.76**	-	0.76**
WMM	-	0.88***	-	0.75**	-	0.60*	-	0.29*
DCM	-	0.85***	-	0.36*	-	0.76**	-	0.76**

Note. * denotes non-significant results with $p \geq .05$, ** denotes $.01 \leq p < .05$, and *** denotes $p < .01$. MWUM: Multiplicative Weights Update Method, RCM: Randomized Correct-rate Method, RSMM: Randomized Simple Majority Method, SMM: Simple Majority Method, WMM: Weighted Majority Method, DCM: Deterministic Correct-rate Method. The Bayesian Information Criterion (BIC) was used to evaluate the performance of stochastic models, with smaller BIC values indicating a better fit. We also calculated the Pearson correlation coefficients between the observed and the predicted proportions of choosing "up." The lowest BIC values and highest Pearson's r among the four conditions is shown in bold.

The results of this research not only contribute to a deeper understanding of human decision-making processes in on-line learning environments, but also highlight the efficacy of MWUM in modeling human behavior. Our findings can be used to enhance the development of decision support systems and expert advice platforms, which can enhance their ability to forecast and facilitate human decision-making. Previous research has demonstrated that humans make decisions in a stochastic manner (Herrnstein, 1961; Vulkan, 2000), and our results serve to reinforce the probabilistic aspect of human decision-making.

Future research could explore the impact of additional factors, such as individual cognitive abilities and preferences, on the adoption of specific on-line decision-making strategies. Moreover, investigating the influence of varying expert profiles or expanding the scope of the task to include more diverse and complex scenarios may provide valuable insights into the generalizability and adaptability of the optimal on-line algorithm in modeling human decision-making across various cognitive domains. We anticipate that ideas from the computer science literature will continue to be useful in exploring how people perform this wider range of on-line optimization problems.

Appendix

In this section, we explore the connections between MWUM and Bayesian inference. We first note that MWUM can be reformulated via the following update rule, maintaining the same optimality guarantee (Arora et al., 2012):

$$w_i^{(t+1)} = w_i^{(t)} e^{-\eta m_i^{(t)}} = w_i^{(1)} e^{-\eta \sum_{\tau=1}^{t} m_i^{(\tau)}} \quad (1)$$

This form of MWUM can be shown to be related to Bayesian inference, treating each expert as a model that could explain the data and trying to infer which model is correct. In the context of binary events, the true model is assumed to generate the observed sequence of "up" and "down" values with an error probability of ε (i.e., there is an ε chance that "up" can be misgenerated as "down" and vice versa). The likelihood function in the Bayesian updating rule is a binomial function, giving

$$p_i^{(t+1)} \propto p_i^{(1)} \prod_{\tau=1}^{t} (1-\varepsilon)^{1-m_i^{(\tau)}} \cdot \varepsilon^{m_i^{(\tau)}} \quad (2)$$

$$= p_i^{(1)} (1-\varepsilon)^t \left(\frac{\varepsilon}{1-\varepsilon}\right)^{\sum_{\tau=1}^{t} m_i^{(\tau)}} \quad (3)$$

where $p_i^{(1)}$ is the prior probability of the i-th model (or expert) being the true data-generating model. Assigning the experts initial weights of 1 is equivalent to assuming a uniform prior. Moreover, the probability of following an expert's advice, acquired by normalizing the weights in Eq.(1), and the posterior probability of selecting a model as the data-generating model, acquired by normalizing Eq.(3), are equivalent when $\eta = \log \frac{1-\varepsilon}{\varepsilon}$. This implies that reducing the error probability $\varepsilon < 0.5$ should increase the step-size parameter in MWUM. The optimal step-size $\eta \approx 0.4$ in MWUM can be construed as around 40% error probability in the Bayesian model. This link also helps connect MWUM with a variety of Bayesian models in cognitive science such as Thompson sampling (Thompson, 1933), which selects an option based on the posterior probability of that model being the best. The stochastic choice procedure in the MWUM is equivalent to selecting the expert with the highest posterior probability, providing a direct equivalence.

Acknowledgments. We thank Princeton Survey Research Center for supplying the online resources necessary for conducting the experiments. We wish to acknowledge the insightful discussions with Huacheng Yu and Matt Weinberg that contributed to this research. This research project and related results were made possible with the support of the NOMIS Foundation.

References

Anderson, J. R. (1990). *The adaptive character of thought*. Psychology Press.

Arora, S., Hazan, E., & Kale, S. (2012). The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, *8*(1), 121–164.

Auer, P., Cesa-Bianchi, N., & Fischer, P. (2002). Finite-time analysis of the multiarmed bandit problem. *Machine Learning*, *47*, 235–256.

Blum, A. (2005). On-line algorithms in machine learning. *Online algorithms: the state of the art*, 306–325.

Cesa-Bianchi, N., & Lugosi, G. (2006). *Prediction, learning, and games*. Cambridge University Press.

Clarkson, K. L. (1988). A las vegas algorithm for linear programming when the dimension is small. In *Proceedings of the 29th Annual Symposium on Foundations of Computer Science* (pp. 452–456).

Foster, D. P., & Vohra, R. (1999). Regret in the on-line decision problem. *Games and Economic Behavior*, *29*(1-2), 7–35.

Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, *55*(1), 119–139.

Freund, Y., & Schapire, R. E. (1999). Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, *29*(1-2), 79–103.

Griffiths, T. L. (2020). Understanding human intelligence through human limitations. *Trends in Cognitive Sciences*, *24*(11), 873–883.

Helmbold, D. P., Schapire, R. E., Singer, Y., & Warmuth, M. K. (1998). On-line portfolio selection using multiplicative updates. *Mathematical Finance*, *8*(4), 325–347.

Herbster, M., & Warmuth, M. K. (1998). Tracking the best expert. *Machine Learning*, *32*(2), 151–178.

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, *4*(3), 267.

Hogarth, R. M. (1977). Methods for aggregating opinions. In *Decision Making and Change in Human Affairs* (pp. 231–255).

Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In *Handbook of the Fundamentals of Financial Decision Making* (pp. 99–127). World Scientific.

Kale, S. (2007). *Efficient algorithms using the multiplicative weights update method*. Princeton University.

Karp, R. M. (1992). On-line algorithms versus off-line algorithms: How much. In *Algorithms, Software, Architecture: Information Processing 92: Proceedings of the IFIP 12th World Computer Congress, Madrid, Spain* (p. 416).

Lai, T. L., Robbins, H., et al. (1985). Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, *6*(1), 4–22.

Littlestone, N., & Warmuth, M. K. (1994). The weighted majority algorithm. *Information and Computation*, *108*(2), 212–261.

Marr, D. (2010). *Vision: A computational investigation into the human representation and processing of visual information*. MIT press.

Rescorla, R. A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. *Classical conditioning, Current research and theory*, *2*, 64–69.

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*(2), 207.

Skinner, B. F. (1966). Contingencies of reinforcement in the design of a culture. *Behavioral Science*, *11*(3), 159–166.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, *25*(3-4), 285–294.

Vovk, V. G. (1995). A game of prediction with expert advice. In *Proceedings of the Eighth Annual Conference on Computational Learning Theory* (pp. 51–60).

Vul, E., Goodman, N., Griffiths, T. L., & Tenenbaum, J. B. (2014). One and done? Optimal decisions from very few samples. *Cognitive Science*, *38*(4), 599–637.

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, *14*(1), 101–118.