COMPUTATIONAL COGNITIVE MODELLING OF NEOCORTEX'S FEATURE: STRATEGIC PLANNING

Mehmet Ali GÜL[1] and Ilham HUSEYINOV[2]

[1,2]Computer Engineering Department, Institute of Natural And Applied Science,
Istanbul Aydin University, Turkey

Corresponding Author:
Mehmet Ali GÜL
Hürriyet Mah. Akce Sok. No:21/4,
Kağıthane, Istanbul – Turkey
Email: mehmetaligul@stu.aydin.edu.tr

Abstract
Today's computational cognitive modeling methods are inspired by various mechanisms of cognition. These models can differ in terms of process detail and input-output detail, and therefore can be implemented at multiple levels. The superficial model of cognition at this level is used as a flexible and radical model. For thousands of years, the human brain has become the environment of the greatest leap of evolution. In particular, the neocortex of the brain, in which the cognitive functions have become evident, has developed advanced opportune methods in the process of performing cognitive functions during the course of its evolution. This tiered and complex order presents us with the flexibility to analyze a more conceptual and meaningful accumulation of information than the first years of the foundation of artificial intelligence. From this inspiration, building a new computational cognitive modelling which is based of neocortex is needed.

Key words: computational cognitive modelling, neocortex, bayesian desicion theory, rational metareasoning

1. Introduction

Models in cognitive and social sciences may be roughly categorized into computational, mathematical, or verbal models. Each model may be viewed as a theory of whatever phenomena it purports to capture. Although each of these types of models has its role to play, the useful and effective way of proposing a new model is given a quantative and implemented model. The reason for this emphasis is that, at least at present, computational modeling appears to be the most promising in many ways, and offers the flexibility and expressive power that no other approaches can match, in terms of providing a variety of modeling techniques and methodologies, or in terms of supporting practical applications [1], [2].

The main point of this article is building an advanced computational cognitive modelling of neocortex toward its feature. The idea of inspiration is from the enhanced human's brain evolution. This will lead to generate effective *Artificial General Intelligence*solutions withimplemented quantative figures which will be shown at outcome of this article.

2. Neocortex's Complex Environments

People are known to use a wide repertoire of different heuristics to make decisions under risk [3]. These strategies include fast-and-frugal heuristics which, as the name suggests, perform very few computations and use only a small subset of the available information [4]. For instance, the lexicographic heuristic (LEX) focuses exclusively on the most probable outcome that distinguishes between the available options and ignores all other possible outcomes. Another fast-and-frugal heuristic that people might sometimes use is Elimination-By-Aspects [5]. Here, we used the deterministic version of EBA described by Payne et al. (1988). This heuristic starts by eliminating options whose payoff for the most probable outcome falls below a certain threshold. If more than one option remains, EBA repeats the elimination process with the second most probable outcome. This process repeats until only one option remains or all outcomes have been processed. After the elimination step EBA chooses one of the remaining outcomes at random. In addition to fast-and-frugal heuristics, people's repertoire also includes more time consuming but potentially more accurate strategies such as the weighted-additive strategy (WADD). WADD first computes each option's expected value, and then chooses the option whose expected value is highest.

In addition to gradually adapting their strategy choices to the structure of the environment [6] people can also flexibly switch their strategy as soon as a different problem is presented. Payne et al. (1988) provided a compelling demonstration of this phenomenon in risky choice: Participants chose between multiple gambles described by their possible payoffs and their respective probabilities. There was a fixed set of possible outcomes that occurred with known probabilities and the gambles differed in the payoffs they assigned to these outcomes. Participants were presented with four types of decision problems that were defined by the presence or absence of a time constraint (15 seconds vs. none) and the dispersion of the outcomes' probabilities (low vs.high); high dispersion means that some outcomes are much more probable than others,whereas low dispersion means that all outcomes are almost equally likely. Ten instances of each of the four problem types were intermixed in random order. The outcomes' payoffs ranged from \$0 to \$9.99, and their values and probabilities were stated numerically. Payne et al. (1988) used process tracing to infer which strategies theirparticipants were using: The payoffs and their probabilities were revealed only when the participant clicked on the corresponding cell of the payoff matrix displayed on the screen, and all mouse clicks were recorded. This allowed Payne and colleagues to measure how often people used the fast-and-frugal heuristics (LEX and EBA) for different types of problems by the percentage of time people spent on the options' payoffs for the most probable outcome. For the expected-value strategy WADD this proportion is only 25%, but for the fast-and-frugal heuristics LEX and EBA it can be up to 100%. The experiment revealed that people adaptively switch decision strategies in the absence of feedback: When the dispersion of outcome probabilities was high, people focused more on the most probable outcome than when all outcomes were almost equally probable. Time pressure also increased people's propensity for such selective and attribute-based processing; see Figure 4. Thus, participants appeared to use fast-and-frugal heuristics more frequently when they had to be fast and when all but one or two outcomes were extremely improbable. This makes sense because the fast-and-frugal heuristics LEX and EBA are fast precisely because they focus on the most predictive attributes instead of integrating all attributes.

3. Simulation of Model

The simulation of the experiment by applying our model to the selection between the ten desicion strategies considered by Payne et al. (1988) including WADD and fast-and-frugal heuristics such as LEX and EBA. To simulate each strategy's execution time counted how many elementary operations (Johnson & Payne, 1985) it would perform on a given problem and assumed that each of them takes one second. This allowed to model to simulate the effect of the time limit on a strategy's performance by having each strategy return its current best guess when it exceeds the time limit (Payne et al., 1988). For the purpose of strategy selection learning, our model represented each decision problem by five simple and easily computed features: the number of possible outcomes, the number of options, the number of inputs per available computation, the highest outcome probability, and the difference between the highest and the lowest payoff. Our model used these features to learn a predictive model of each strategy's relative reward;

$$r_{\mathrm{rel}}(s;o) = \frac{V(s(D),o)}{\max_a V(a,o)},$$

where $s(D)$ is the gamble that strategy s chooses in decision problem D, $V(c, o)$ is the payoff of choice c if the outcome is o, and the denominator is the highest payoff the agent could have achieved given that the outcome was o. To choose a strategy, the predicted relative reward $r$rel is translated into the predicted absolute reward r by the transformation;

$$\hat{r} = \min\{r_{min} + (r_{max} - r_{min}) \cdot \hat{r}_{rel}, r_{max}\},$$

where r_{min} and r_{max} are the smallest and the largest possible payoff of the current gamble respectively. The model then integrates the predicted absolute reward and the predicted time cost into a prediction of the strategy's VOC according to Equation 2 and chooses the strategy with the highest VOC as usual. The priors on all feature weights of the score and execution time models were standard normal distributions. The simulation assumed that people knew their opportunity cost and did not have to learn it from experience. Rather than requiring the model to learn the time cost as outlined above, theopportunity cost was set to $7 per hour and normalized by the maximum payoff ($10) to make it commensurable with the normalized rewards.

To compare people's strategy choices to rational metareasoning, we performed 1000 simulations of people's strategy choices in this experiment. In each simulation, we modeled people's prior knowledge about risky choice strategies by letting our model learn from ten randomly generated instances of each of the 144 types of decision problems considered by Payne et al. (1988). We then applied rational metareasoning with the learned model of the strategies' execution time and expected reward to a simulation of Experiment 1 from Payne et al. (1988). On each simulated trial, we randomly picked one of the four instances and generated the payoffs and outcome probabilities according to the problem type: Outcome distributions with low dispersion were generated by sampling outcome probabilities independently from the standard uniform distribution and dividing them by their sum. Outcome distributions with high dispersion were generated by sampling the outcome probabilities sequentially such that the second largest probability was at most 25% of the largest one, the third largest probability was at most 25% of the second largest one, and so on. Since the participants in this experiment received no feedback, our simulation assumed no learning during the experiment.

To evaluate our theory against alternative hypotheses, we also ran 1000 simulations according to SCADS. To evaluate our theory against alternative hypotheses, we also ran 1000 simulations according to the SCADS model. We did not evaluate SSL or RELACS because these theories do not distinguish different kinds of problems and hence cannot account for the phenomena observed by Payne et al. (1988).

The SCADS model was equipped with nine categories (time pressure, no time pressure, many options (> 3), few options (≤ 3), many possible outcomes (> 3), few possible outcomes (≤ 3), high stakes (range of payoffs ≥ 50% of highest payoff), lowstakes (range of payoffs ≤ 10% of highest payoff), and non-compensatory (largest outcome probability > 0.5)). As before, we considered three instances of SCADS whose reward functions were either the relative payoff, the relative payoff minus the opportunity cost of the strategy's execution time, or the reward rate. The SCADS model's categoryspecific and global strategy-reward associations were initialized with strengths equivalent to one observation per strategy.

We found that rational metareasoning correctly predicted that time-pressure and probability dispersion increase people's propensity to use the fast-and-frugal heuristics LEX and EBA; see Figure 1. Time pressure increased the predicted frequency of fast,attribute-based processing by 29.69% ($t(1998) = 9.70$, $p < 10^{-15}$), and high dispersionof the outcome probabilities increased the predicted frequency of fast, attribute-based processing by 44.11% ($t(1998) = 14.41$, $p < 10^{-15}$). Furthermore, their strategy choices only change in response to reward or punishment but the experiment provided neither. The SCADS model can choose strategies adaptively in principle, but in our simulations its strategy choices were dominated by the global, problem-independent associations. Consequently, our SCADS models always converged onto a single strategy during the training phase and continued to do so in the test trials. Hence, the predicted effects of time pressure (−0.1 to 0%, all $p \geq 0.4955$) and dispersion (0% to 0.05%, all $p \geq 0.4978$) were not significantly different from zero. In conclusion, rational metareasoning can account for adaptive flexibility in decision-making under risk but SSL, RELACS, and SCADS cannot.These results suggest that rational metareasoningcan capture the adaptive flexibility of people's strategy choices not only for behavioral strategies that manipulate external representations but also for cognitive strategies that operate on internal representations.

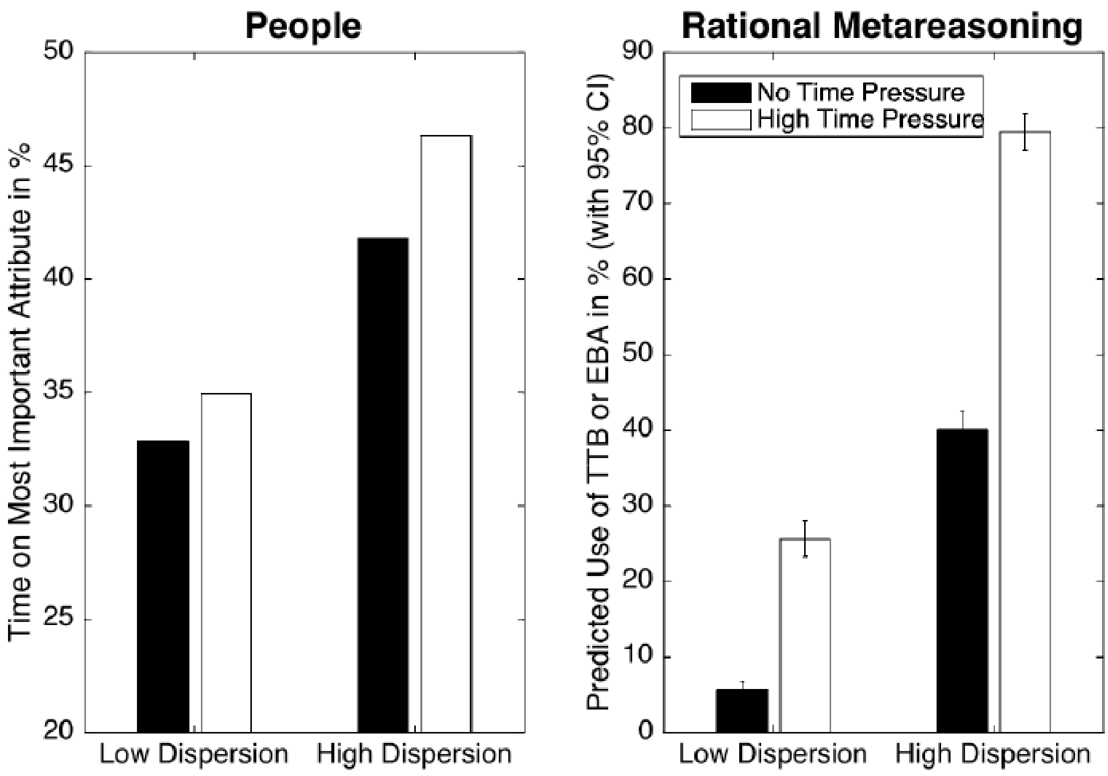


Figure 1: Rational metareasoning predicts the increase in selective attribute-based processing with dispersion and time pressure observed by Payne et al. (1988).

4. Results

To test our hypothesis that people learn to deliberate less, we classified the participants' response patterns into three categories: The response strategy on a trial was categorized as random choice if the participant chose one of the gambles without inspecting any of the outcomes. If the participant chose "No thanks!" without inspecting the outcomes, then the response strategy was classified as disengaged. Finally, if the participant clicked on at least one of the outcome boxes, then the response was categorized as engaged. We measured our participants' performance on the task by three metrics: engagement, reward rate, and adaptive randomness. Our model predicted that participants' reward rate and adaptive randomness would increase significantly from the pretest to the posttest while their engagement decreases.

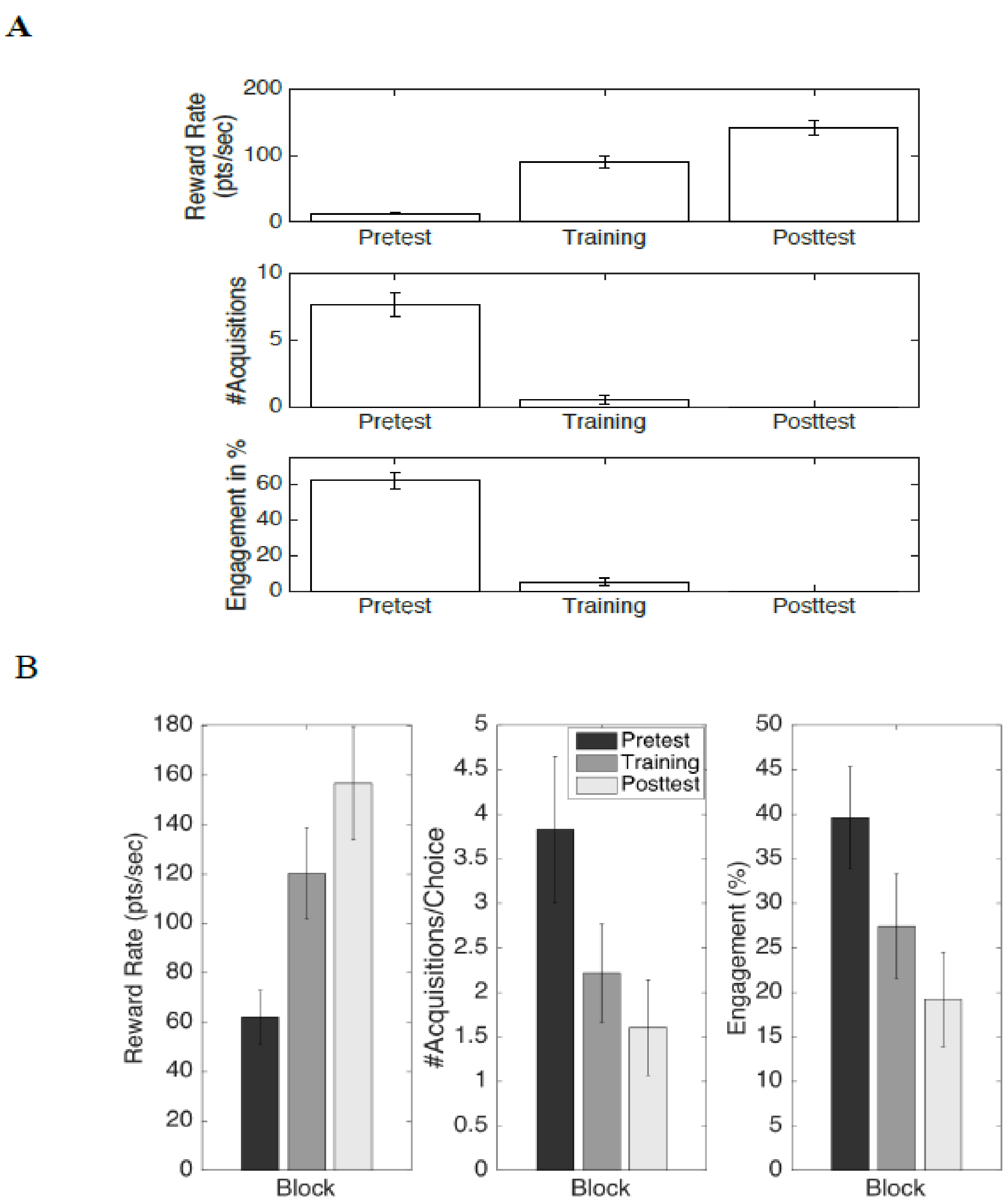


Figure 2: Experiment 2: Learning when not to engage in effortful decision-making. A: Predictions of rational metareasoning for Experiment 2. B: The empirical findings of Experiment 2 confirmed the three qualitative model predictions.

As shown in Figure 2A, our rational model predicted that participants should learn to decide more quickly and thereby win increasingly more points per second by engaging in deliberation less often and acquiring fewer pieces of information. Since the simulated decision-maker estimates its reward rate by Bayesian inference as defined above, it gradually realizes that its opportunity cost is very high. In addition, the simulated decision-maker learns that deliberate strategies are slow, and that the random strategy performs about as well as deliberation when all outcomes are similar. Hence, the simulated decision-maker eventually learns to avoid deliberating, to skip problems with negative payoffs, and to apply the random strategy when all outcomes are great.

As shown in Figure 6B, we found that the learning induced changes in our participants' strategy choices were consistent with our theory's predictions. There was a significant increase in the participants' average reward rate ($t(99)$ = 9.98, $p < 10^{-15}$; Cohen's d = 1.00) as they learned to process less information ($t(99) = -4.80$, $p<10^{-5}$; Cohen's $d = -0.48$) and their engagement decreased significantly ($t(98) = -7.89$, $p < 10^{-11}$; Cohen's $d = -0.79$). Even though participants acquired increasingly less information, their average reward per decision did not change signific significantly from the first block to the last block ($t(98)$ = 0.69, p=0.49; Cohen's d = 0.07).

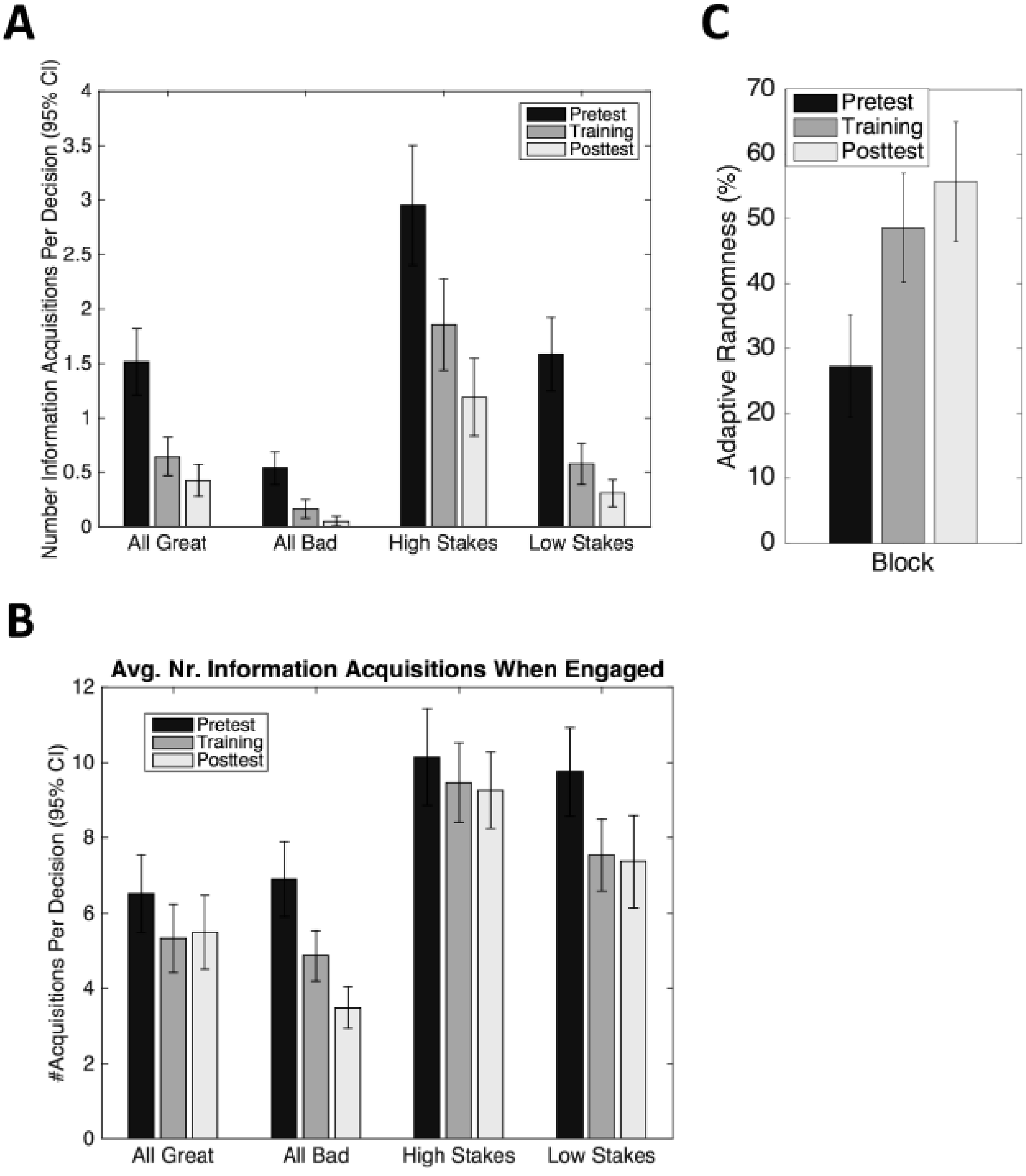
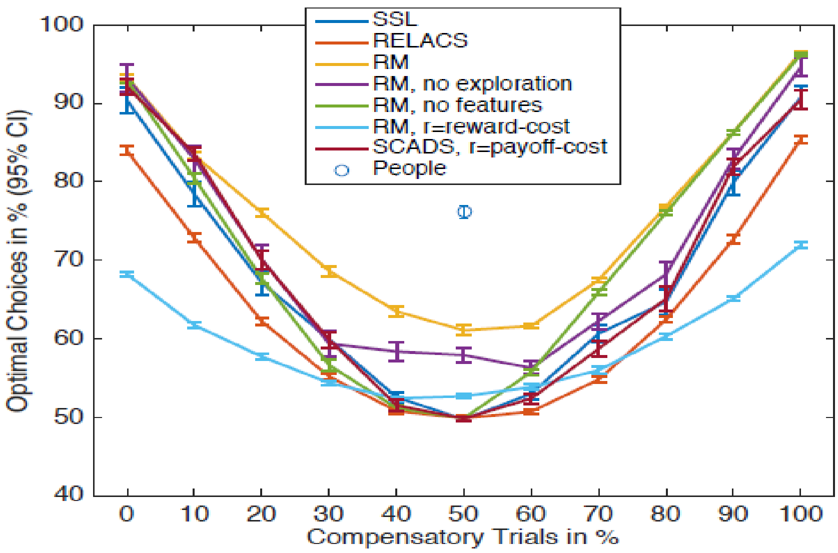


Figure 3: Adaptive disengagement in Experiment 2. A: Average number of information acquisitions by block and problem type. B: Number of information acquisitions when engaged. C: Adaptive randomness increased as participants learned to apply the random choice strategy more often to problems where all outcomes were great and less often to other problems.

In summary, Experiment 2 placed participants in an environment where maximizing the reward rate required choosing without deliberation, and the participants learned to reap increasingly higher reward rates by acquiring increasingly fewer pieces of information, choosing at random when all outcomes were great and to skipping all other problems. There was also a trend towards learning to prioritize the most probable outcome. All of these effects are consistent with the hypothesis that people learn to make increasingly more rational use of their finite time and computational resources.

A



B

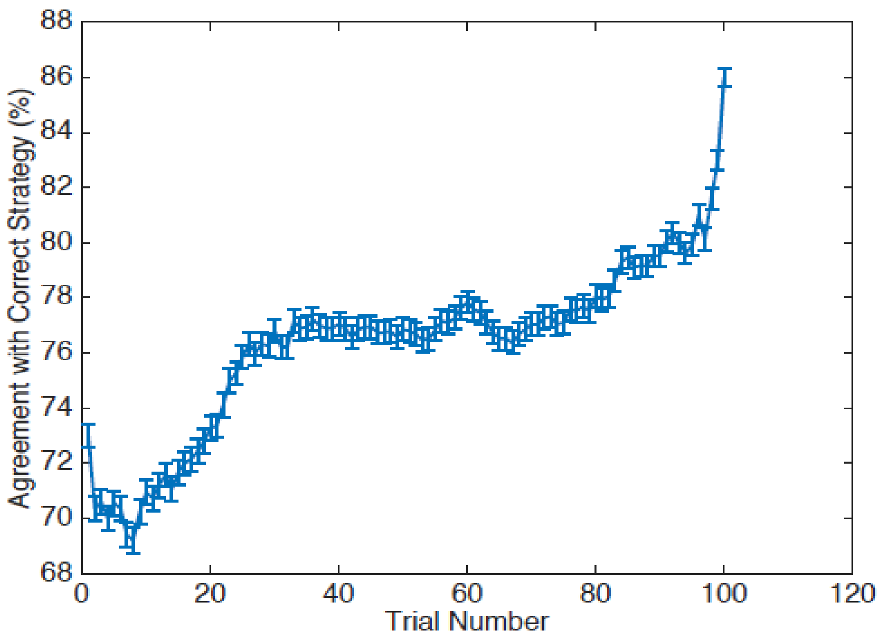


Figure 4: Model predictions and findings of Experiment 4. A: People and rational metareasoning perform significantly above chance in heterogeneous environments but context-free strategy selection mechanisms do not. B: People's performance increased with experience. The trial-by-trial frequencies were smoothed by a moving average over 20 trials. The error bars enclose 95% confidence intervals.

5. Conclusion

The experiments presented in this section confirmed the predictions of our resource-rational theory of strategy selection learning: The first experiment showed that people learn to think less when they think too much. The second experiment showed that people learn to think more when they think too little. Thirdly, we showed that people learn to adapt not only how much they think but also how they think to the structure of the environment. Finally, Experiment 2 demonstrated that adaptive flexibility also increases with learning, and this enables people's strategy choices to exploit the structure of individual problems. Most importantly, in all four cases, the underlying learning mechanisms made people's strategy choices increasingly more resource-rational. Hence, the empirical evidence presented in this section supports our hypothesis that the human brain is equipped with learning mechanisms that make it more resource-rational over time. Even though people may not be resource-rational when they first enter a new environment, the way in which they process information appears to converge to the rational use of their finite time and bounded computational resources.

6. Future Directions

Future work should extend the proposed model to capture additional aspects of human cognition. One such extension could be a more realistic model of the cost of strategy execution which captures that some strategies are more effortful than others. This could be achieved by modeling how much cognitive resources, such as working memory, each strategy consumes at each point in time. With this extension, the total cost of executing a strategy could be derived by adding up the opportunity costs of its consumed resources over the time course of its execution.

References

[1] Pew, R. W., Baron, S. (1978). *The components of an information processing theory of skilled performance based on an optimal control perspective.* In Steinbach, G. E. (Ed.), Information processing in motor control and learning (pp. 71–78). New York: Academic Press.

[2] Ritter, Shadbolt, Elliman, Young, Gobet, &Baxter(2003). *Techniques for Modeling Human Performance in Synthetic Environments: A Supplementary Review.* Including a Model of Situation Awareness and Rapid Decision Making. (pp. 60).Human Systems Information Analysis Center Wright-Patterson Air Force Base, Ohio.

[3] John W. Payne, James R. Bettman, Eric J. Johnson (1993). *The Adaptive Decision Maker*. Cognitive effort and decision strategies (pp. 75). Cambridge University Press.

[4] Gerd Gigerenzer and Wolfgang Gaissmaier (2011). *Heuristic Decision Making*.Journal Vol. 62:451-482. Center for Adaptive Behavior and Cognition, Max Planck Institute for Human Development, 14195 Berlin, Germany.

[5] Shafir, E.B., Simonson, I., & Tversky, A. (1993). Reasonbased choice. *Cognition*, Review 49, 11-36.

[6] Rieskamp J. (2008). *The importance of learning when making inferences*. Judgment and Decision Making, Vol. 3, No. 3(pp. 261–277). Max Planck Institute for Human Development.