High-order feature-based mixture models of classification learning predict individual learning curves and enable personalized teaching

Yarden Cohen and Elad Schneidman[1]

Department of Neurobiology, Weizmann Institute of Science, Rehovot 76100, Israel

Edited by Ranulfo Romo, Universidad Nacional Autonoma de Mexico, Mexico City, D.F., Mexico, and approved November 13, 2012 (received for review July 7, 2012)

Pattern classification learning tasks are commonly used to explore learning strategies in human subjects. The universal and individual traits of learning such tasks reflect our cognitive abilities and have been of interest both psychophysically and clinically. From a computational perspective, these tasks are hard, because the number of patterns and rules one could consider even in simple cases is exponentially large. Thus, when we learn to classify we must use simplifying assumptions and generalize. Studies of human behavior in probabilistic learning tasks have focused on rules in which pattern cues are independent, and also described individual behavior in terms of simple, single-cue, feature-based models. Here, we conducted psychophysical experiments in which people learned to classify binary sequences according to deterministic rules of different complexity, including high-order, multicue-dependent rules. We show that human performance on such tasks is very diverse, but that a class of reinforcement learning-like models that use a mixture of features captures individual learning behavior surprisingly well. These models reflect the important role of subjects' priors, and their reliance on high-order features even when learning a low-order rule. Further, we show that these models predict future individual answers to a high degree of accuracy. We then use these models to build personally optimized teaching sessions and boost learning.

inference | information | maximum entropy

We regularly learn to classify sensory stimuli into novel categories, often using an impoverished sampling of the stimulus space and the underlying rule: even a "simple" task of classifying patterns of n bits into two categories, requires an implicit mapping of the 2^n possible patterns, which means there are 2^{2^n} potential deterministic classification rules. It is clear then that when we learn to classify, we cannot simply explore the space of rules and patterns, but instead must rely on simplifying assumptions.

Analysis of human learning of deterministic classification rules has focused on modeling of the average behavior of subjects (1–3), and explored the effect of rule complexity on the average level of success (4). Learning to classify according to a probabilistic rule is inherently ambiguous, and so studies of such tasks have focused on simpler rules than those used in deterministic classification. For example, the weather prediction (WP) task (5) requires learning probabilistic associations between multiple cues and a label, where each cue carries independent information about the correct label. Analysis of the learning strategy of individual subjects in this task has compared single-cue or single feature-based strategies, and the possibility of switching between such strategies (6, 7). Associative or Bayesian learning models that rely on simple stimulus features were used to describe the diversity of individual learning dynamics that subjects exhibited and compare between subjects (8), and reflected differences between healthy subjects and patients (9, 10). However, these models were mostly evaluated in terms of their ability to describe subjects' performance, rather than cross-validated predictive power. Second, the complexity of the rules studied was limited (i.e., cues carried independent information). Third, these tasks often relied on a strong bias in the presentation rate of different patterns, which changes the available information for subjects.

To characterize learning behavior of complex rules at the individual level, we used a psychophysical task of classifying binary visual patterns into two abstract classes, with no bias in the set of presented patterns. Our task involved only deterministic rules, but included rules of different complexity, focusing on high-order dependencies between elements in the pattern. We extended ideas of feature-based learning (6, 11–13) to present a reinforcement learning-like family of maximum entropy-based models to describe how individuals learn different high-order classification rules. We found that these models capture individual behavior to a high degree of accuracy, and can also predict individual behavior. We then used such models, which were fitted to individuals during a learning session, to pick the optimal samples to present them to help them learn the rule faster.

Results

To characterize individual classification learning in terms of accurate quantitative models, we presented 41 healthy subjects with a psychophysical task in which they had to classify patterns of black and white squares, into two abstract categories labeled "red" and "blue" (Fig. 1*A*). The correct label of each pattern was determined by a deterministic rule—a fact that was unknown to the subjects. In each session, patterns of size $n = 4$ or 5 squares were presented one at a time, and the correct label was shown after the subject classified the pattern. To enable quantitative analysis and comparison between subjects, all subjects were presented with the same order of patterns; moreover, each block of 16 examples (for $n = 4$, or 32 patterns in case $n = 5$) contained all possible patterns in a "frozen" random order. Given the huge number of deterministic rules (65,536 for $n = 4$, and 9×10^9 for $n = 5$), we chose six different balanced rules (equal number of red- and blue-labeled patterns), but of different complexity. Denoting the patterns as $\vec{x}$, where $x_i = \pm 1$, and the label as $y = \pm 1$, the label of each pattern was determined according to single, pairwise, or triple-wise dependencies between the squares in the pattern, or according to "holistic" features of the whole pattern (one-bit, two-bit, three-bit, majority, symmetry, and middle symmetry rules; Fig. 1). Each subject performed four sessions on the same day, with a different rule in each session.

The diversity of learning dynamics among subjects was wide, ranging from no learning, to incremental learning, to abrupt transition to full success. Fig. 1*B* shows the fraction of correct answers as a function of time (over a running-average window) for 10 different subjects learning the same 1-bit rule ($y = x_3$). Similar

Author contributions: Y.C. and E.S. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: elad.schneidman@weizmann.ac.il.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1211606110/-/DCSupplemental.

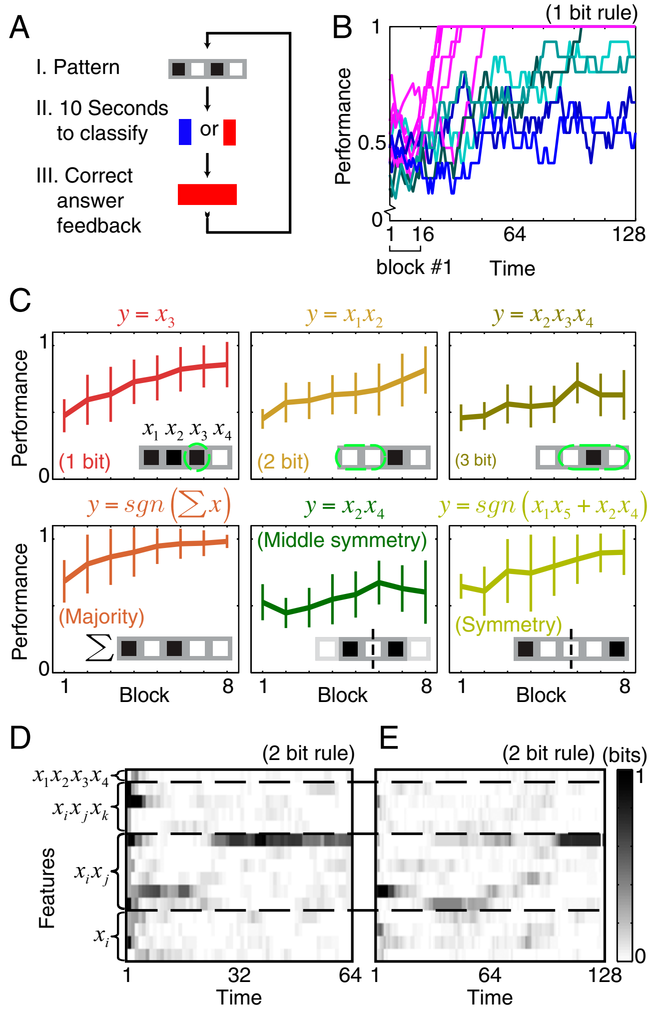

Fig. 1. Pattern classification learning task and diversity of individual performance. (*A*) Setup of the psychophysical task. In each step of an experimental session, a pattern of black and white squares was shown on a screen, and subjects had 10 s to classify it as blue or red. Then the correct label of the pattern was presented, and a new pattern was presented. (*B*) Diversity of individual learning curves. The running average performance of 10 subjects classifying patterns according to a one-bit rule, with the same order of patterns. For clarity, we color-coded learning curves of subjects who demonstrated no clear improvement during the 128 patterns (blue), gradual learning (cyan), and abrupt learning (magenta). (*C*) Average learning curves for different rules used. Each panel shows the average learning curve over n = 12–39 subjects for one rule used, $y = y(\vec{x})$. Performance mean and SD over subjects were estimated in blocks of 16 steps. (*D*) The mutual information between the subject's choices and different features of the pattern, estimated over running windows of 20 steps. This subject had a clear, single feature-based classification rule, which was switched from one pairwise feature to another. (*E*) Same as in *D*, but for a subject that did not have a clear single feature-based strategy.

diversity of learning curves was seen for the other rules we tested. Fig. 1*C* shows the average learning curve over all subjects for the different rules, which reflect the average difficulty of each rule. The large SDs over subjects result from the individual differences between them; in particular, for each rule there was at least one subject that failed to learn it, and no subject succeeded in learning all rules. The performance for the majority rule demonstrates the strong effect of the prior that subjects had, as 2 of 16 subjects learned the rule without any mistakes, and another two made a single mistake (12). The marked differences between the population learning curves of the different rules reflect that memorization without generalization can be ruled out as the sole mechanism of learning. In *SI Text, Ruling out Simple Memorization as the Strategy That Subjects Use*, we show that both gradual memorization and pattern specific memorization can also be ruled out as pure strategies.

Intuitively, one might assume that subjects seek distinctive features in the patterns according to which they classify, as has been studied in similar tasks (1, 4, 6). We therefore quantified the relation between a set of different features of the patterns, $f_i(\vec{x})$, and subjects' answers. Fig. 1 *D* and *E* show the mutual information between subjects' choices and pattern features, $I(f_i;\ answers)$ as a function of time for different one-, two-, three-, and four-bit features (this set is a complete basis from which one could linearly construct any other feature). In a few cases, subjects indeed relied on single features (Fig. 1*D*), but in most cases, they were using a different strategy, which could be described as mixing of features, or a very fast switching between features (Fig. 1*E*).

We therefore used a mixture model of these features, which are a basis that can span any rule, to explore individual behavior and characterize the effect of subjects' priors on learning (Fig. 2*A*). Using Bayes' rule we can represent any probabilistic classifier of pattern x into label y, in terms of "internal models" that a subject has for each category c (the probability that pattern x belongs to category c) as a weighted mixture of features of x,

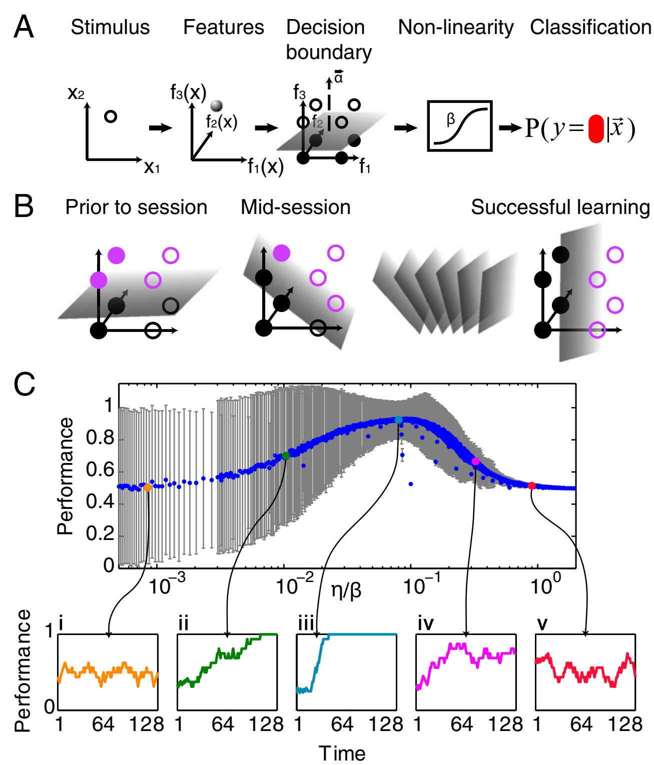

Fig. 2. Mixture model description and theoretical analysis. (*A*) Graphic representation of the mixture of features-based model. The classifier is a hyperplane in the space of features, f, of the pattern x. The features mixture coefficients, $\vec{\alpha}$, define the decision boundary plane in the features space. We extract a global scaling parameter, β, which serves as the model's confidence. The logit form of Eq. **2** results from Bayes' rule. (*B*) Graphic representation of learning dynamics. At the start of the experimental session, subjects' decisions can be explained in terms of a hyperplane in the high dimensional space of pattern features. During the learning process, this hyperplane is shifted and rotated in the features space, getting it closer to the correct rule. (*C*) Model performance as a function of η:β. The mean model performance (blue dots) and SD (gray bars) for a one-bit rule are computed using a random set of initial conditions for each value of η:β. (*i*–*v*) Typical model learning curves.

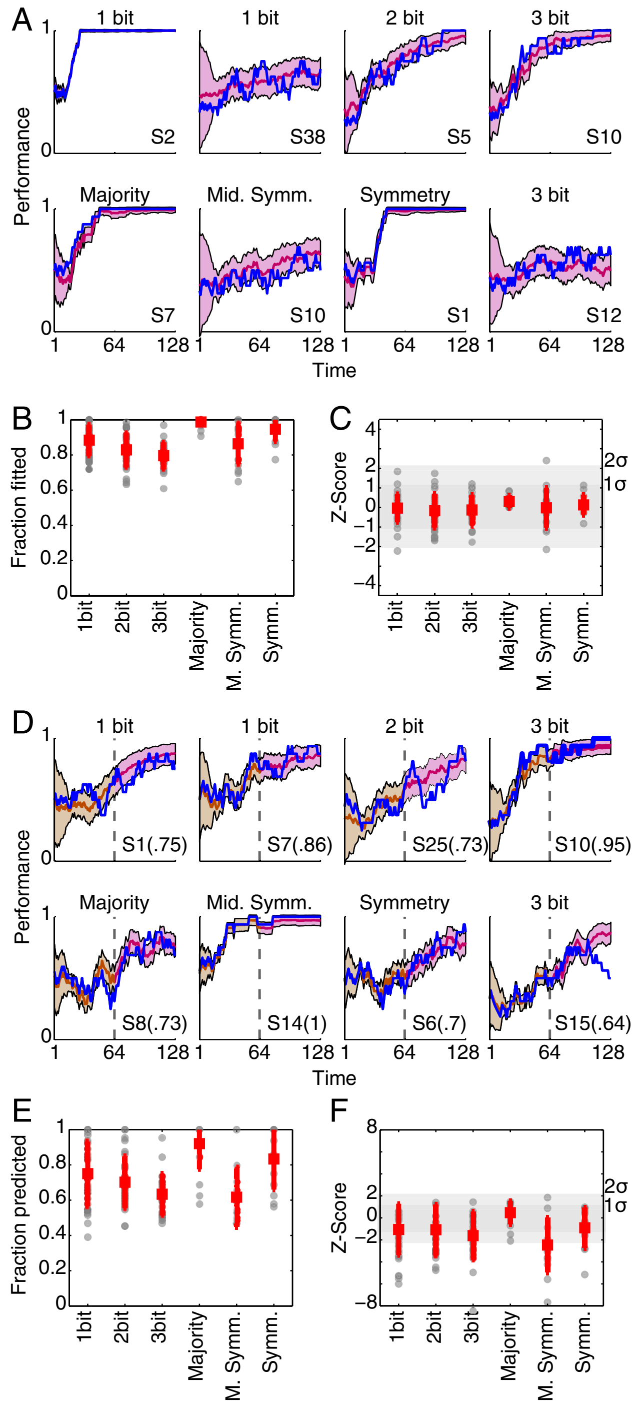

Fig. 3. Mixture of features models accurately describe individual performance and predict subjects' future answers. (*A*) Examples of models' fit to individual subjects. The subjects' learning curves (blue lines) compared with the fitted machine expected learning curves (purple lines) and its SD (pink area) (*B*) Fraction of each individual's answers that agree with those of the fitted model (gray dots) was computed using the most likely answers. The red markers and bars show the population mean and SD. (*C*) For each individual model we estimate how typical the empirical answer sequence was by dividing the agreement between the model and answers by the expected SD of the model's answers. These Z-score values are estimated for each individual session and subject (gray dots), and the red markers and bars are the population's mean and SD. (*D*) Examples of models' prediction of individual subjects. The subjects' learning curves (blue lines) compared with the fitted machine expected learning curves (brown lines) and its SD (light brown area), and the predicted learning curve (purple lines) and its SD (pink area). (*E*) Fraction of answers predicted correctly in the second half of the trial, based on a model that was fitted in the first half. Gray dots mark the fraction of agreement for each session, and the red markers and bars are the population's mean and SD. (*F*) As in *C*, Z-score values for predicted answers.

$$P(\vec{x}|y=c)=\frac{1}{Z_c}exp\left(\beta\sum_i \alpha_i^c\cdot f_i(\vec{x})\right), \qquad [1]$$

where $f_i(\vec{x})$ are features of $\vec{x}$; α_i are the weights given for each feature, normalized such that the norm of $\vec{\alpha}$ is 1, β quantifies the certainty or inverse "temperature", and Z_c is a normalization factor, or partition function. This form of exponential functions is commonly used in machine learning, statistics, psychophysics, and neuroscience (14–16). We note that for a specific choice of features, this is the maximum-entropy distribution of $\vec{x}$, given the average value of that set of features in each category, $\langle f_i(\vec{x})\rangle_c$. The classifier $p(y|x)$ is then given by

$$P(y=1|\vec{x})=\frac{1}{1+\frac{Z_1}{Z_{-1}}exp\left(\beta\left(\vec{\alpha}\cdot\vec{f}(\vec{x})+\gamma\right)\right)}, \qquad [2]$$

where γ is given by the category prior, $\frac{P(y=-1)}{P(y=1)}=e^{\gamma}$, $\vec{\alpha}=\vec{\alpha}^{-1}-\vec{\alpha}^{1}$, and Z_1 is the partition function of label 1 in Eq. **1**. This model can describe any classifier and learning dynamics in terms of arbitrary changes to each of the α_is at each time-step.

Importantly, to find a compact and useful representation of individual learning in terms of a subject's prior and simple set of learning parameters, we used a very restricted version of learning the model described by Eqs. **1** and **2**. Instead of changing the weight of each feature independently, we consider models in which learning occurs by changing all α_is in a coupled way, according to a single learning rate and a gradient-based rule: given a labeled example, the change in weights is given by

$$\Delta\alpha_i^c=\eta\frac{\partial \log P(y|\vec{x})}{\partial \alpha_i^c}=-\eta\left(1-P(y|\vec{x})\right)\cdot\left(\langle f_i(\vec{x})\rangle_{p(\vec{x}|c)}-f_i(\vec{x})\right)\cdot y\cdot c, \qquad [3]$$

where η is a single static learning rate, $\langle f_i\rangle_{p(\vec{x}|c)}$ is the expected value of feature f_i over the model distribution given by Eq. **1**, and γ has a similar learning rule as in Eq. **3**. Thus, the model starts with a prior set of weights for each of the features α_i at $t=0$, and then updates the weights according to a learning rule that can be interpreted as a product of a standard reinforcement learning term (13, 17) and a gradient-based term. Thus, the parameters of our models are just the prior at $t=0$ and two static parameters: the certainty β and learning rate η (Fig. 2*B*). We note that this is a generalized form of the models used by Speekenbrink and Shanks (11) for the WP task, but here they include a specific prior term and high-order features.

These models can exhibit a wide range of learning dynamics, which we mapped by simulating the performance of models with different values of η and β on different rules, with a large set of diverse priors. Fig. 2*C* shows the success rate of the simulated models in learning the one-bit rule (Fig. 1*C*) as a function of the ratio $\eta{:}\beta$. We found that for low $\eta{:}\beta$, the prior dominates the model's behavior. For large $\eta{:}\beta$, the prior has no effective role, because the learning rule is overfitting to the last seen sample, resulting in poor generalization. Optimal learning in this case would occur at $\eta{:}\beta\simeq 0.1$, which would balance the prior with appropriate learning rates. (We show in *SI Text, Dependence of Learning in the Mixture of Features Model on* $\eta{:}\beta$, that the ratio $\eta{:}\beta$ dominates the learning dynamics of these models).

We thus fitted the subjects' learning sessions with our mixture of features model, and found that the detailed structure of their individual learning curves can be approximated to a high degree

of accuracy. Specifically, for each subject and session, we used a heuristic approach to find the best prior and combination of β and η values that would maximize the likelihood of the observed sequence of answers (*Methods*). Fig. 3*A* shows examples of the high similarity between the empirical learning curve of subjects and the expected learning curve of the model that was fit to them. Because the models are probabilistic, we estimated the variance of the model's learning curve, and found that the models are almost as similar to the observed behavior as one could expect. Fig. 3*B* shows the overlap in answers between subjects and the models that were fit to them, for all rules. Averaging over all rules and subjects, the accuracy was 87%; the average overlap for the different rules ranged from 80% to 98%. To further validate the accuracy of the model, we asked, how typical is the observed sequence of answers according to the model we fitted? Fig. 3*C* shows the overlap between the model and the empirical answers, normalized by the SD of the model's predictions. The Z-score values show that the deviation between models and the data were within the probabilistic variance expected from the model.

Though our model relies on a spanning set of features, it is possible that subjects do not use or have access to all these features. We asked then how removing features from the basis set of the model would affect the accuracy and predictive power of the model. We found that mixture models that used only single-bit or pairwise features were significantly less accurate than models that used higher-order features in fitting subjects' behavior (Fig. 4), almost regardless of rule complexity. A more general learning rule, which evaluated the change in $\vec{\alpha}$ based on feature values over a window of several examples $\langle f_i(\vec{x})\rangle_W$, rather than just the last one, did not significantly improve our results. Neither did other reinforcement learning models and neural network-based ones (*SI Text, Comparisons of the Mixture of Features Model to Other Models and Effect of the Length of the Learning Window on the Mixture of Features Model*).

We asked then whether our models can predict individual behavior. Fig. 3*D* shows examples of the agreement between the subjects' learning curves on the second half of the session, and the most likely learning curves predicted by the model that was fit just to the first half of that session. On average, we predicted ~75% of the future answers (Fig. 3*E*). Fig. 3*F* shows Z-score measures of the prediction accuracy, reflecting that though these models are missing some part of the subjects' learning dynamics, they are very informative of individual subjects.

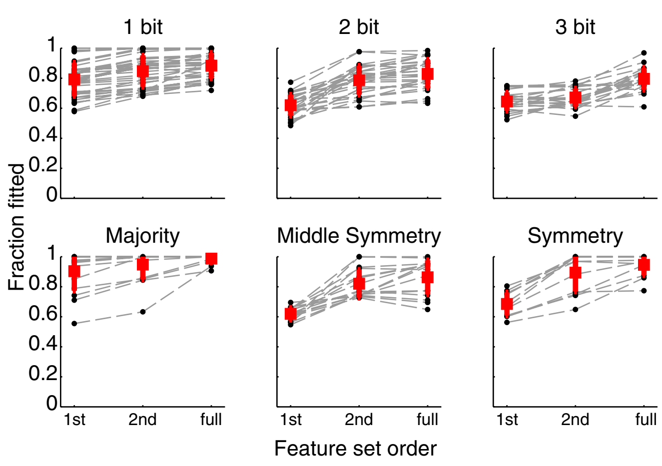

Fig. 4. High-order features set improves model fitting for all of the rules. Every session was fitted with models, whose features set contained only first order features, first and second order features or a full set of features (*x*-axis). Each panel groups the sessions of a different rule and dashed lines connect the fits of the same session with models of different features' sets. The black dots' *y*-axis value represents the fraction of fitted answers.

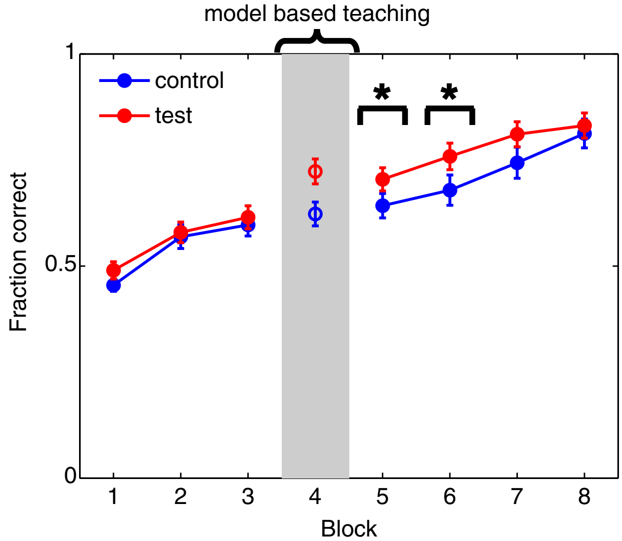

Fig. 5. Model-based personalized teaching improves subjects' learning curves. The fraction of correct answers in blocks of 16 steps is used to compare the test group (red) and the control group (blue): the test group had samples shown to them in block 4, according to the model's prediction of the best samples to show each person. The control group had a random set of samples presented to them in block 4. Markers show population mean, and error bars are SEM. Stars mark blocks in which there was a significant effect ($P < 0.04$, unpaired t test).

Having a predictive model of a subject's future behavior suggests that we could use the model to estimate which patterns would be most useful for the subject in terms of learning the rule. We thus repeated the same experiment, but now presented to the subjects a personalized sequence of patterns in the fourth block of samples. To determine the individual set of patterns we should show each subject, we used the first three blocks (where all subjects had the exact same patterns shown to them) to fit a mixture model to each subject (online). If we now assume that the model is an accurate predictor of the subject's behavior, we can simulate the effect of any sequence of patterns in the fourth block on learning. We then sought the optimal set of patterns to present the subject in that fourth block ("personalized training block") that would result in the most accurate classifier at the end of that block, according to the model. In *SI Text, Pattern Specific Performance Before Personalized Training Block, and the Order of Patterns in That Block*, we show that the optimal sequences used in the fourth block did not take a trivial form and also contained patterns that were classified well in the first three blocks. To evaluate the effect of the individually chosen group of patterns, all subjects had the exact same patterns shown to them in blocks 5–8. We found that both the test group and our control group, which were presented with a random set of patterns in the fourth block, improved during the session. However, the test group showed a significant improvement in the two blocks following the intervention (unpaired t test, $P < 0.04$; Fig. 5). The control group closed this gap later in the session. In *SI Text, Optimality of the Personalized Training Sequence: Comparison with Alternative Models*, we show that choosing the sequences according to other models is expected to give worse learning improvement. Thus, we succeeded in boosting individual performance by personalized, model-based teaching. We note that this approach resembles ideas from optimal sampling and active learning in statistics and machine learning (18, 19, *), which suggest future analysis of optimal exploration and exploitation strategies in learning.

*Roy N, McCallum A (June 28–July 1, 2001) Toward optimal active learning through sampling estimation of error reduction. *Proceedings of the 18th International Conference on Machine Learning* (Morgan Kaufmann, San Francisco), pp 441–448.

NEUROSCIENCE
PSYCHOLOGICAL AND COGNITIVE SCIENCES

Discussion

We have shown that in a pattern classification task, human performance can be described to a high degree of accuracy by a probabilistic model that dynamically adapts the weight subjects give to different features of the patterns. These models rely on a prior and two static parameters, and were accurate enough to enable personally tailored teaching, by picking for individual subjects the best patterns to show them to help them learn.

Our results reflect the important role of the prior that a subject has in such learning tasks (12), and that these priors already rely on high-order correlations in the patterns that the subjects classify. This finding suggests that subjects' history and experience are instrumental in shaping their learning.

We note that it is possible that more-detailed models, in particular using adaptive learning rates (η) or subject's certainty (β), could result in better fit for the subjects' answers and predictive power of the model. However, we submit that the power of the current approach is in its relative simplicity.

The mathematical nature of our models suggests that learning may be seen as a dynamic weighting of "experts" † by combining, linearly, prototypes of the stimulus or exemplar representation of the stimulus space (2, 20, 21) (*SI Text, Mixture Model as a Simple Case of Prototype or Exemplar Representation*). Because the classification rule could rely on an arbitrary feature of the patterns—and, in particular, some of the rules we studied were nonlinear functions of simple features of the patterns (in contrast to refs. 5 and 22)—our model included in the mixture a set of features that allowed describing any rule, i.e., a spanning set of features. This features set could correspond to neural correlates of decision-making (23, 24), probabilistic inference (13, 25), and learning and strategy shifts (26, 27) that were observed in single-unit recordings in primate and mammalian cortex. Seeking neural correlates of the model presented here would be of particular interest in light of the characterization of the role of memory systems involved in WP (28, 29) and other learning and decision-making tasks (30, 31), and theoretical models of incremental learning through spike timing-dependent plasticity (32, 33).

Methods

Subjects. A total of 78 healthy adults (34 male) between the ages of 22 and 42 performed four learning sessions, taking short breaks between sessions, for a total duration of ∼1 h. The experimental setup was explained to all subjects before the first session, and they went through one brief training session. The detailed nature of the complexity of classification rules was not discussed with the subjects. Subjects signed a formal participation agreement according to Helsinki protocol TLV-0287-09 approved by the Weizmann Institute's institutional review board. Subjects were rewarded for participation, regardless of their performance. *SI Text, Instructions to Subjects in the Pattern Classification Task*, contains the instructions that were given to the subjects.

Experimental Setup. Stimulus presentation, feedback, and recoding of responses was done using MatLab (MathWorks) and Psychtoolbox (34) on a standard desktop computer (Intel-based operating system: Windows XP, 4 GB RAM). Subjects indicated their choices using a standard computer mouse.

Behavior analysis. Subjects' answers were binary and so their performance was measured via (*i*) learning curves, using a moving average window of 16 steps to smooth the stepwise performance, and (*ii*) block averages, averaging the performances in consecutive blocks of 16 steps.

The mutual information between the subject's choice, r, and the value of a parity feature, $f_J(\vec{x}) = \prod_{j\in J} x_j$, is estimated from the empirical estimation of the rates, $P(r, f_J)$, as

$$I(r; f_J) = \sum_{r=\pm1} \sum_{f_J=\pm1} P(r, f_J) \cdot \log_2\left(\frac{P(r, f_J)}{P(r)P(f_J)}\right). \quad [4]$$

Modeling. The model's learning dynamics is given by a gradient ascent step (Eq. **3**) as described in the text, followed by renormalizing $|\vec{\alpha}^c| = 1$.

For a given set of model coefficients, Θ, including a prior, $\vec{\alpha}^{-1}$, $\vec{\alpha}^{1}$ at $t = 0$ ($\gamma_{t=0} = 0$), a certainty, β, and a learning rate, η, the learning algorithm produces a set of decision probabilities, $P(y_t = 1|\vec{x}_t; \Theta)$, for the patterns sequence $\{\vec{x}_t\}_{t=1}^{128}$. We then use these for model-fitting and prediction of future answers.

Model-fitting. We treat the subject's sequence of answers, $\{r_t\}_{t=1}^{128}$, as a realization of the binomial independent random variables, $p_t \sim Binom(1, P(1|x_t; \Theta))$, and fit each session of each subject with a model that maximizes the log-likelihood:

$$L(\Theta) = \sum_{t=1}^{128} \log P(p_t = r_t|x_t). \quad [5]$$

Because this function is not concave, we heuristically looked for the optimal point, Θ^*, by maximizing the likelihood through a combination of a genetic algorithm and simulated annealing.

Prediction of subjects' future behavior based on their individually fitted model. For prediction we fit the model based only on the first 64 answers in a session. We then estimated the agreements between the model and the subject as the fraction of subject answers that match the model's most likely answers, $a = \left\langle r_t = sgn\left(P(1|x_t; \Theta) - \frac{1}{2}\right)\right\rangle_{t\in T}$. When estimated, the expected agreement if the subject's answers were indeed a realization of the model's decision probabilities, $\mu = \langle \max(P(1|x_t; \Theta), 1 - P(1|x_t; \Theta))\rangle_{t\in T}$, and its SD, $\sigma = \frac{1}{|T|}\left[\sum_{t\in T} P(1|x_t; \Theta) \cdot (1 - P(1|x_t; \Theta))\right]^{1/2}$; we used these to estimate the Z-score, $z = \frac{a-\mu}{\sigma}$, as a measurement of difference between the subject and the model.

Personalized Teaching. Subjects had to learn a two-bit rule (with $n = 4$). All of the subjects were presented with the same patterns, except for the fourth block, in which they were shown a sequence of patterns that was optimized for them individually. To fit the model online, we prepared a set of 80 million candidate models, covering a large region of parameter space, for which we stored the decision probabilities for learning steps 1–48 and the most helpful sequence for steps 49–64. The most helpful sequences were found using a *Glauber* dynamics search (35) that aimed to bring the models to the closest point to the target rule. While subjects performed steps 1–48, their answers were transmitted, using transmission control protocol/internet protocol, to another computer that calculated the best fit among the 80 million models and transmitted back the patterns sequence for the fourth block. The entire transmission duration was always less than the waiting time between samples (1 s), and so did not affect the session progress.

ACKNOWLEDGMENTS. We thank Rony Paz and Peter Dayan for valuable suggestions and comments. This work was supported by the Peter and Patricia Gruber Foundation, the Israel Science Foundation, and the Clore Center for Biological Physics (E.S.).

†Roth S, Black MJ (June 20–25, 2005) Fields of experts: A framework for learning image priors. *Computer Vision and Pattern Recognition, 2005. IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (IEEE, San Diego), Vol 2, pp 860–867.

1. Shepard RN, Hovland CI, Jenkins HM (1961) Learning and memorization of classifications. *Psychol Monogr* 75(13):1–42.
2. Kruschke JK (1992) ALCOVE: An exemplar-based connectionist model of category learning. *Psychol Rev* 99(1):22–44.
3. Nosofsky RM, Kruschke JK, McKinley SC (1992) Combining exemplar-based category representations and connectionist learning rules. *J Exp Psychol Learn Mem Cogn* 18(2):211–233.
4. Feldman J (2000) Minimization of Boolean complexity in human concept learning. *Nature* 407(6804):630–633.
5. Knowlton BJ, Squire LR, Gluck MA (1994) Probabilistic classification learning in amnesia. *Learn Mem* 1(2):106–120.
6. Gluck MA, Shohamy D, Myers C (2002) How do people solve the "weather prediction" task?: Individual variability in strategies for probabilistic category learning. *Learn Mem* 9(6):408–418.
7. Lagnado DA, Newell BR, Kahan S, Shanks DR (2006) Insight and strategy in multiple-cue learning. *J Exp Psychol Gen* 135(2):162–183.
8. Shanks D, Speekenbrink M (2007) Through the looking glass: A dynamic lens model approach to multiple-cue learning. Economic Learning and Social Evolution Working

Papers 240 (Economic and Social Research Council Centre for Economic Learning and Social Evolution, London).
9. Speekenbrink M, Lagnado DA, Wilkinson L, Jahanshahi M, Shanks DR (2010) Models of probabilistic category learning in Parkinson's disease: Strategy use and the effects of L-dopa. *J Math Psychol* 54(1):123–136.
10. Meeter M, Radics G, Myers CE, Gluck MA, Hopkins RO (2008) Probabilistic categorization: How do normal participants and amnesic patients do it? *Neurosci Biobehav Rev* 32(2): 237–248.
11. Speekenbrink M, Shanks DR (2008) Through the looking glass: A dynamic lens model approach to multiple cue learning. *The Probabilistic Mind*, eds Chater N, Oaksford M (Oxford Univ Press, London), pp 409–429.
12. Goodman ND, Tenenbaum JB, Feldman J, Griffiths TL (2008) A rational analysis of rule-based concept learning. *Cogn Sci* 32(1):108–154.
13. Soltani A, Wang X-J (2010) Synaptic computation underlying probabilistic inference. *Nat Neurosci* 13(1):112–119.
14. Pillow JW, et al. (2008) Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454(7207):995–999.
15. Beck JM, et al. (2008) Probabilistic population codes for Bayesian decision making. *Neuron* 60(6):1142–1152.
16. Ackley DH, Hinton GE, Sejnowski TJ (1985) A learning algorithm for Boltzmann machines. *Cogn Sci* 9(1):147–169.
17. Rescorla R, Wagner A (1972) *Classical Conditioning II: Current Research and Theory*, eds Black A, Prokasy W (Appleton-Century-Crofts, New York), pp 64–99.
18. Cohn DA, Ghahramani Z, Jordan MI (1996) Active learning with statistical models. *J Artif Intell Res* 4(1):129–145.
19. Freund Y, Seung HS, Shamir E, Tishby N (1997) Selective sampling using the query by committee algorithm. *Mach Learn* 28(2):133–168.
20. Estes WK (1986) Memory storage and retrieval processes in category learning. *J Exp Psychol Gen* 115(2):155–174.
21. Posner MI, Keele SW (1968) On the genesis of abstract ideas. *J Exp Psychol* 77(3): 353–363.
22. Sigala N, Logothetis NK (2002) Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* 415(6869):318–320.
23. Yang T, Shadlen MN (2007) Probabilistic reasoning by neurons. *Nature* 447(7148): 1075–1080.
24. Kepecs A, Uchida N, Zariwala HA, Mainen ZF (2008) Neural correlates, computation and behavioural impact of decision confidence. *Nature* 455(7210):227–231.
25. Soltani A, Lee D, Wang XJ (2006) Neural mechanism for stochastic behaviour during a competitive game. *Neural Netw* 19(8):1075–1090.
26. Law CT, Gold JI (2009) Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nat Neurosci* 12(5):655–663.
27. Nassar MR, Wilson RC, Heasly B, Gold JI (2010) An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J Neurosci* 30(37):12366–12378.
28. Poldrack RA, et al. (2001) Interactive memory systems in the human brain. *Nature* 414(6863):546–550.
29. Foerde K, Knowlton BJ, Poldrack RA (2006) Modulation of competing memory systems by distraction. *Proc Natl Acad Sci USA* 103(31):11778–11783.
30. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ (2006) Cortical substrates for exploratory decisions in humans. *Nature* 441(7095):876–879.
31. Dayan P, Daw ND (2008) Decision theory, reinforcement learning, and the brain. *Cogn Affect Behav Neurosci* 8(4):429–453.
32. Farries MA, Fairhall AL (2007) Reinforcement learning with modulated spike timing dependent synaptic plasticity. *J Neurophysiol* 98(6):3648–3665.
33. Florian RV (2007) Reinforcement learning through modulation of spike-timing-dependent synaptic plasticity. *Neural Comput* 19(6):1468–1502.
34. Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10(4):433–436.
35. Glauber RJ (1963) Time-dependent statistics of the Ising model. *J Math Phys* 4(2): 294–307.