Consistency in Brain Activation Predicts Success in Transfer

Qiong Zhang[1,2], **John R. Anderson**[2,4], **Robert E. Kass**[1,3,4]
qiongz@andrew.cmu.edu ja+@cmu.edu kass@stat.cmu.edu
[1]Machine Learning Department, [2]Psychology Department, [3]Statistics Department, [4]Center for the Neural Basis of Cognition
Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213 USA

Abstract

Recent brain imaging studies have provided new insight into how students are able to extend their previous problem solving skills to new but similar problems. It is still unclear, however, what the basis is of individual differences in their success at transfer. In this study, 75 subjects had been trained to solve a set of mathematical problems before they were put into the fMRI scanner, where they were challenged to solve modified versions of familiar problems. A hidden semi-Markov model identified the sequential structure of thought when solving the problems. Analyzing the patterns of brain activity over the sequence of states identified by the model, we observed that subjects who showed consistent brain patterns performed better. This consistency refers to both how consistently subjects respond to different problems (within-subject consistency), and how brain responses of a given subject deviate from the population average (between-subjects consistency). Early within-subject consistency is particularly predictive of later performance in the experiment.

Keywords: Transfer of learning; Mathematical problem-solving; Individual differences; Hidden semi-Markov model

Introduction

While sometimes the learners are skilled only at a task being taught, in many cases they are expected to transfer what they have learned to new situations. It is the fundamental assumption in education that what is learned will apply in similar but possibly different situations (Leberman, McDonald, & Doyle, 2006). A specific setting that has raised considerable interest is the transfer of learning during mathematical problem solving. For decades, various methods have been used to shed light on the underlying thought processes during complex mathematical problem solving. Such efforts include protocol analysis and eye movements (Epelboim & Suppes, 2001). Functional Magnetic Resonance Imaging (fMRI) has become a powerful instrument to collect vast quantities of data on brain activity. We have developed a new procedure that combines multi-voxel pattern analysis (MVPA) (Norman, Polyn, Detre, & Haxby, 2006) and hidden Markov model (HMM) algorithms (Rabiner, 1989) to better analyze the temporal patterns in the brain (Anderson, Betts, Ferris, & Fincham, 2010). This method is particularly effective in mathematical problem solving, where there is a rich mixture of perceptual, cognitive and motor activities each with distinct temporal characteristics. Given the great variability in problem-solving duration of such problems, we have used a hidden semi-Markov model to identify the sequential structure and durations of the problem solving process.

In recent years, brain imaging studies have informed us about the neural basis and mechanisms underlying transfer of learning (Ischebeck, Zamarian, Schocke, & Delazer, 2009) (Anderson & Fincham, 2014b). However, there remains an open question: what are the sources of individual differences in successful transfer? This question is the focus of our study. In our experiments, subjects were trained in a mathematical problem-solving task before they were scanned. In the fMRI sections of the experiments, they encountered both Regular problems that were like those they had solved, as well as novel Exception problems that required subjects to devise modifications or partial replacements to their learned procedure.

Methods

Experiments

Our research uses *pyramid problems*, which involve a mathematical relationship denoted by a dollar symbol as the operator, e.g., 4\$3 = X. Here, 4 is the base, which is also the first term in an additive sequence; 3 is the height, which indicates the number of terms to add in a descending manner, e.g. 4\$3 = 4 + 3 + 2. This is an example of a "Regular" problem. There are "Exception" problems that either have unusual numbers for operands, e.g. 4\$-3=X, or have unusual algorithm such as having X as one of the operands, e.g. X\$4 = 30. Solving Exception problems involves modifying the procedure for Regular problems. The exact problems and procedural details are described in the original report (Anderson & Fincham, 2014a). fMRI data were collected from 40 adults (ages 19-35), and 35 children (ages 12-14). Subjects practiced solving a large number of Regular problems outside the scanner, and were tested on the second day in an fMRI scanner with a mixture of Exception problems and Regular problems. Each subject solved six blocks of problems with each block consisting of two Regular problems and nine Exception problems. Images were acquired using gradient echo-echo planar image (EPI) acquisition on a 3T Verio, then motion corrected and co-registered. Blood-oxygenation-level dependent (BOLD) signal response is calculated as the percent change from a linear baseline defined by the first scan. This is deconvolved with a hemodynamic response function to produce an estimate of the underlying activity signal using a Wiener filter (Glover, 1999). The hemodynamic function is the difference of two gamma bases (Friston, 2007).

Dimensionality Reduction

To reduce the dimensionality of the data and to accommodate variations in anatomy over subjects, the original image voxels (3.125 x 3.125 x 3.2 mm) were aggregated into larger 2x2x1 voxels within the same slice. Removing those showing extreme values resulted in 8365 voxels. To further reduce the dimensionality of the 8365 voxels, principal component analysis was performed with 67% of the variance captured by the first 20 components that we eventually worked with. In addition, we examined the BOLD brain activations of the 12 key regions of interest (ROIs) averaged over the left side and the right side of the brain.

Discovering Mental States

Hidden Markov models (HMM) have been successfully used in modeling and analyzing complex behavioral and neurophysiological data (Obermaier, Guger, Neuper, & Pfurtscheller, 2001). They make inference of unobserved parameters possible while taking into account the probabilistic nature of behavior and brain activity. We conceive the subjects as going through a sequence of mental states in the pyramid experiments. The discrete mental states are hidden in the sense that we only observe the brain activity, not the mental states themselves. The effectiveness of HMMs in modeling such an experiment has already been demonstrated in a previous study, where three ground-truth stages of fixation, problem solving, and feedback could be recovered accurately on a single-trial basis (Anderson & Fincham, 2014a). Four states were identified as the optimal number in the previous study with the same experiment through leave-one-out cross-validation (LOOCV). They were characterized as an encoding stage, a planning stage, a solving stage, and a responding stage, given their brain signatures (Anderson & Fincham, 2014a).

Model Specification. A specific extension of HMM, a hidden semi-Markov model (HSMM), is used to model explicitly the state duration as a gamma distribution that provides a realistic characterization of response latencies (Wainer & Messick, 1983). The state duration is discretized to the nearest scan. The probability of spending m scans in state i given the length of each scan being 2 seconds is:

$$P(m; v_i, a_i) = \int_{2m-1}^{2m+1} g(t; v_i, a_i)\,\mathrm{d}t,$$

where $g(t; v_i, a_i)$ is the gamma distribution with shape parameter v_i and scale parameter a_i. The fMRI activity considered in the model are the first 20 components obtained from principal component analysis over all scans in the experiment. They are further normalized to have mean 0 and standard deviation 1. The brain activity of the k^{th} PCA component for each state i is modeled as a normal distribution $N(x; \mu_{ik}, 1)$ with mean μ_{ik} and standard deviation 1. The probability of observing a set of PCA components $F_j = \{f_{j1}, f_{j2}, .., f_{j20}\}$ for a particular scan j, at state i, is calculated as:

$$P(F_j; M_i) = \prod_{k=1}^{20} N(f_{jk}; \mu_{ik}, 1),$$

where $M_i = \{\mu_{i1}, \mu_{i2}, .., \mu_{i20}\}$. We have a left-right HSMM that encodes a linear sequence of four mental states. The implementation that we adapted is from the software developed by Yu and Kobayashi (Yu & Kobayashi, 2006).

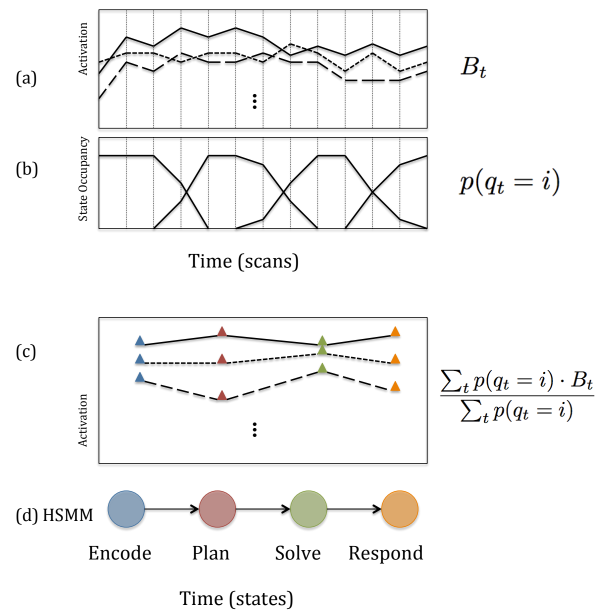

Figure 1: An illustration of how to obtain state-specific brain activations for each of the 4 states in a single trial for different brain regions, i = 1,2,3,4. (a) shows the brain activations for each scan for representative brain regions in one trial. (b) shows the state occupancies estimated from the HSMM of that trial. (c) shows the state-specific brain activations calculated from weighted sums for each region in (a). (d) shows the linear structure of the HSMM model. Each vertical line in (a) and (b) represents an fMRI scan.

Trial-alignment By States. Modeling with an HSMM not only uncovers sequential stages of problem solving, but also serves as an effective way to align fMRI data on a trial-by-trial basis. Conventionally, trials would be aligned to a fixed time point, either the stimulus or the response. However, because different processes will occur at the same time on different trials, the average signal can be uninformative (Gibbons & Stahl, 2007; Velazquez & Wennberg, 2009). In our analysis, after fitting all trials to the HSMM, state occupancies for each scan $p(q_t = i)$ can be obtained. They are the probabilities that each scan belongs to each of the four states, where q_t is the state at time t, and i = 1,2,3,4. As is illustrated in Figure 1, brain activations for each scan are converted to brain activations for each state through a sum weighted by the state occupancies. These state-specific brain activations will be used throughout our analysis.

Approach

Our study is an exploratory analysis that attempts to answer the question: what contributes to individual differences in transfer-of-learning performance? Past research has found that the extent to which subjects are able to relate the current task to the appropriate past experiences contributes to success in transfer tasks (Ross, 1984). One hypothesis is that this facilitates taking a consistent approach to the problems. Our current study will examine whether there is evidence for the importance of consistency at the brain level. We start our investigation of this question by characterizing what is meant by brain consistency. This consistency can be understood from two perspectives. It can be measured as the agreement in brain activation patterns amongst different problems in an experiment for a given subject (within-subject consistency). It can also be measured as the agreement in brain activation patterns from a given subject with the rest of the subjects (between-subjects consistency). These brain consistencies are first explored on a set of pre-defined regions of interest (ROIs), and then tested at the whole-brain level on a set of selected brain voxels.

Results

The Pre-defined Brain Regions

A set of pre-defined brain regions of interest (ROIs) have been observed to play an important role in the complex problem solving in our laboratory. Eight of these are long-standing regions in ACT-R cognitive modeling (Anderson et al., 2004): the fusiform gyrus (Talairach coordinates of the ROI center: -42, 60, -9), the secondary auditory area (-46, 22, 9), the caudate nucleus (-14, -10, 7), the lateral inferior prefrontal cortex (-43, -23, 24), the posterior parietal cortex (-23, 63, 40), the anterior cingulate cortex (-6, -10, 39), the manual portion of the motor-sensory region (-42, 19, 50), and the facial portion of the motor-sensory region (-43, 13, 33). Recently, Anderson and Fincham defined a new metacognitive module in ACT-R that involves the rostrolateral prefrontal cortex (-33, -47, 8) (Anderson & Fincham, 2014b). We will also use an additional three regions identified by Dehaene as important to mathematical cognition: the horizontal segment of the intraparietal sulcus (-34, 49, 45), the angular gyrus (-42, 65, 37), and the posterior superior parietal lobule (-19, 68, 54) (Dehaene, Piazza, Pinel, & Cohen, 2003).

We investigated how the consistency in a subject's activation in this set of 12 ROIs predicted subject performance. Subject performance is measured as the proportion of problems one solves correctly during the entire experiment, rather than response time (10.88s per problem on average). This consistency can be measured within each subject as the averaged correlation between every pair of problems across the 12 ROIs (within-subject consistency), reflecting how consistently one subject responds to different problems. It can also be measured as the correlation between the subject mean and the population mean across the 12 ROIs (between-subjects consistency), reflecting how consistently one subject solves problems compared with the rest of the subjects. To avoid any confounding of our measures with accuracy we only looked at correlations among correct problems. Although adults (78.71%) perform better than children (62.16%), the results do not substantially change if analyzed separately. A dataset of this size, with two populations pooled together, provides a basis for application of the HMM state-discovery procedures.

Better subjects exhibit a higher level of brain consistencies. The 75 subjects can be divided into two groups based on their performance: 38 better subjects (85.28%) and 37 weaker subjects (56.31%). Analysis focuses on the early stage (the first 10), the late stage (the last 10), or the entire experiment (all). Figure 2 shows the within-subject consistencies for the two groups of subjects calculated separately for each state during the early stage and the late stage of the experiment. Analysis of variance (ANOVA) was performed on these brain consistencies where the factors were group (better vs. weaker), period (first 10 versus last 10), and state (encode, plan, solve, respond). There are significant effects of all three factors (group: $F(1,73) = 18.45, p < .0001$), period: $F(1,73) = 24.69, p < .0001$), and state: $F(3,219) = 52.09, p < .0001$), such that brain consistencies are greater for better subject, later states, and the first 10 problems. There is a marginal interaction between group and period ($F(1,73) = 3.41, p = .069$), such that the difference between better and weaker subjects decreases from the first 10 problems to a smaller difference for the last 10 problems (Figure 2).

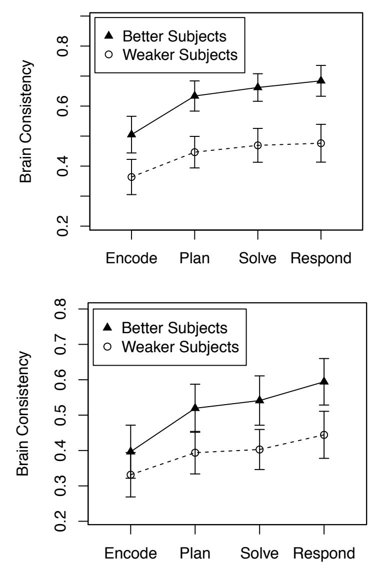

Figure 2: Brain consistency (within-subject) for the better subjects and the weaker subjects during the first 10 correctly solved problems (Top) and the last 10 correctly solved problems (Bottom). Error bars show the 95% confidence interval of the population means.

To further quantify the relation at the level of individual difference, a second level of correlation is carried out between within-subject brain consistency (average correlation) and subject overall performance. The larger the values are, the more correlated the measure of consistency with the subject performance. Brain activations in this analysis are obtained with respect to each of the four states - Encode, Plan, Solve and Respond. In our analysis, Pearson's Correlation Coefficient of larger than 0.227 is considered significant ($p < 0.05$) under two-tailed probabilities with a sample of size of 75 (df = 73). As is shown in Table 1, there is significant correlation between within-subject brain consistency and subject overall performance for most of the states. This effect is stronger for the first 10 correctly solved problems than the last 10 correctly solved problems. The same effect is also observed over the entire experiment, when analyzing all correctly solved problems.

Table 1: Pearson correlations between subject performance and the *within-subject* brain consistency among either the first 10 correctly solved trials, the last 10 correctly solved trials, or for all correctly solved trials, for each of the four states. p-value is indicated below each value of correlation. Significant ($p < 0.05$, two-tailed) correlations are in bold.

States	Encode	Plan	Solve	Respond
First 10	**0.420**	**0.582**	**0.568**	**0.512**
	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Last 10	0.163	**0.313**	**0.346**	**0.351**
	0.16	0.0062	0.0024	0.002
All	**0.285**	**0.439**	**0.461**	**0.441**
	0.013	< 0.0001	< 0.0001	< 0.0001

There is also significant correlation between between-subjects consistency and subject overall performance as shown in Table 2 analyzed in a similar manner as that of the within-subject consistency. Analysis of variance (ANOVA) was performed on these brain consistencies where the factors were group (better vs. weaker), period (first 10 versus last 10), and state (encode, plan, solve, respond). Again there are significant effects of all three factors (group: $F(1,73) = 13.58, p < .0001$), period: $F(1,73) = 23.06, p < .0001$), and state: $F(3,219) = 40.43, p < .0001$).

Table 2: Pearson correlations between subject performance and *between-subjects* brain consistency among either the first 10 correctly solved trials, the last 10 correctly solved trials, or for all correctly solved trials, for each of the four states. p-value is indicated below each value of correlation. Significant ($p < 0.05$, two-tailed) correlations are in bold.

States	Encode	Plan	Solve	Respond
First 10	**0.230**	**0.314**	**0.314**	**0.317**
	0.047	0.0061	0.0061	0.0056
Last 10	**0.267**	**0.377**	**0.353**	0.164
	0.021	0.00086	0.0019	0.16
All	**0.327**	**0.382**	**0.389**	**0.274**
	0.0041	0.00072	0.00056	0.017

The effect of between-subjects consistency reveals that better subjects deviate less from the global average. This observation relates to a study of the Space Fortress task where the similarity to a global definition of mental stages obtained from all the subjects can classify the mental stages of a particular scan for individual subjects (Anderson, Bothell, Fincham, & Moon, 2014). We will focus for the rest of the study on the examination of the within-subject consistency whose effect is stronger as illustrated in Table 1, and argue that this effect arises specifically from the transfer task.

Extending the problem solving procedure from a familiar problem to a new one is a challenging task. Subjects were trained on the Regular problems before going into the scanner and having to transfer this knowledge to Exception problems. At the beginning of the last century, Thorndike and Woodworth proposed that the amount of transfer depends on how many shared elements there are between the learned tasks and the transfer tasks, which is now widely known as the theory of Identical Elements (Thorndike & Woodworth, 1901). This theory was later refined by Gick and Holyoak, when they brought out the concept of perceived similarity (Gick & Holyoak, 1987). Perceived similarity depends on not only the objective number of shared elements, but also the knowledge or expertise of the person performing the transfer task. The more a subject can relate the current transfer tasks to the past learned tasks, and perceive them similarly, the more transfer will take place (Gick & Holyoak, 1987).

In our pyramid experiment, although every subject was presented with the same set of problems, these problems might be perceived very differently. It is likely that the consistency at the neural level in our correlation analysis reflects how similarly the set of new problems appear to the subjects, compared with the trained problems. The more one finds the new and modified problems similar to the trained ones, the more one is able to use knowledge about having solved the trained problems. The perceived similarity has also been identified in the analogy literature as whether an abstract shared structure is learned, with various areas in prefrontal cortex (PFC) postulated to be responsible for enabling analogy (Geake & Hansen, 2005; Green, Fugelsang, Kraemer, Shamosh, & Dunbar, 2006; Bunge, Wendelken, Badre, & Wagner, 2005). Instead of examining only the BOLD response in a single region at a single time, we looked into pairwise correlations among problem pairs across a network of brain regions at different points of problem solving, to explicitly define the perceived similarity at the neural level. Aly& Turke-Brown have recently used a similar correlation approach to show that better memory performance is predicted with attention-stabilized activation patterns in the hippocampus (Aly & Turk-Browne, 2015).

Early consistency is predictive of the overall performance. To see how well we could use this within-subject brain consistency for prediction we used a Leave-one-subject-out cross validation (LOOCV) procedure. In particular, the performance of an unseen subject is predicted by weights trained from the rest of the subjects using multiple regression analysis with Least Squares Fitting. The four independent variables (predictors) here are the measures of consistency among the first 10 correctly solved problems for each of the four states. The dependent variable (what we are predicting) is the subject performance, measured as the proportion of problems solved correctly during the entire experiment. As is observed in Figure 3, there is a considerable match between the predicted performance using LOOCV and the actual accuracy of the subjects, with the correlation of the two being 0.548, and the mean squared error (MSE) being 0.0393. Thus it can be concluded that the within-subject brain consistency among the first 10 correctly solved problems is not only an indicator but also an effective predictor of the overall performance of an unseen subiect.

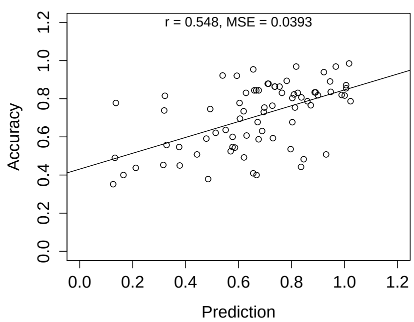

Figure 3: Leave-one-subject-out cross validation (LOOCV) performance prediction of 75 subjects compared with the actual subject performance, using the weights of multiple regression analysis obtained from the rest of the subjects. A total of 4 predictors are 4 within-subject brain consistencies measured across 12 ROIs among the first 10 correctly solved problems for each of the 4 states. (r = .548, MSE = .0393).

The same analysis is also carried out using the late brain consistency instead of the early consistency, which gives 0.346 instead of 0.548 as the correlation between the predicted performance and the actual accuracy, and 0.0799 as MSE. A model using all 8 predictors (both early and late) is significantly better than a model using only later predictors (F (4, 66) = 5.6679, p = .00055), but not than a model using only early predictors (F (4, 66) = .7869, p = .538). It can be concluded that including the early consistency will improve the prediction using the late consistency, but not vice versa.

The previous analysis indicated that late consistency is significantly less predictive than early consistency. Also, as Figure 2 reveals, late consistency is lower. However, if consistency indicates expertise, with more training, why it would decrease? We suspect that initially successful subjects are taking a common approach to the problems. But as time goes on, subjects start to develop more diverse and problem-specific strategies. Thus subjects may start displaying more diverse brain patterns as a result of their more diverse solution strategies. It is interesting to note in this regard that the better subjects show a larger decrease in brain consistency from the early stage to the late stage of the experiment (Figure 2).

The Selected Brain Voxels

Previously, brain consistency was measured as correlation across the pre-defined 12 ROIs. However, our explanation of the consistency effect suggests the effects should not be limited to just these predefined regions, but show up in all regions that are engaged by the task. Therefore, we apply a whole brain analysis identifying all engaged regions and testing whether consistency across these regions predicts success on the task. We identify engaged regions as those brain voxels whose average activation on the first 10 correctly solved problems significantly correlated with the subject performance ($r = .2272, p < 0.05$, two-tailed). Note that this selection method does not imply anything about consistency of activation across voxels, which is the property that we will only use to predict performance. To examine quantitatively how well the measure of consistency can predict the subject performance, a similar procedure of leave-one-subject-out cross validation (LOOCV) is used. First, we select the regions to correlate by determining which are the engaged voxels for the other 74 subjects. Then for each state and each subject, we calculate the within-subject brain consistency – the mean correlation of activations across selected voxels over every pair of problems during the early stage of the experiment. Then we regress the four brain consistencies of the 74 subjects against their overall performance. The coefficients of this regression are used to predict the performance of the 75th subject, given the subject's four brain consistencies corresponding to each state.

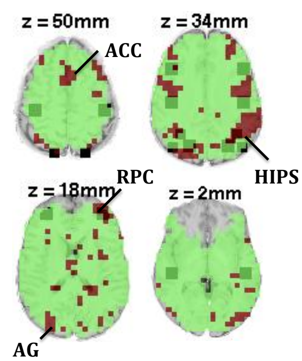

Figure 4: Visualization of the selected brain voxels in red whose averaged activations over the first 10 correctly solved problems are significantly (p≤ 0.05) correlated with the subject overall performance, for a representative state - Encode, and on one group of 74 subjects. The rest of brain voxels are colored light green. Predefined ROIs that overlap with the selected regions are noted.

Figure 4 illustrates the brain voxels selected for one group of 74 subjects. It also shows that these regions overlap with some of the 12 pre-defined regions like the angular gyrus (AG), the horizontal segment of the intraparietal sulcus (HIPS), anterior cingulate cortex (ACC), and the rostrolateral prefrontal cortex (RPC). The exact number of the selected voxels differs for each subset of 74 subjects, but is around one tenth of the total number of voxels. As is observed in Figure 5, there is a considerable match between the predicted performance using LOOCV and the actual accuracy of the subjects, with the correlation of the two being 0.640 and MSE being 0.029. LOOCV with selected brain voxels further improves the performance prediction compared with previously using only the 12 ROIs with correlation of 0.548 and MSE of 0.0393.

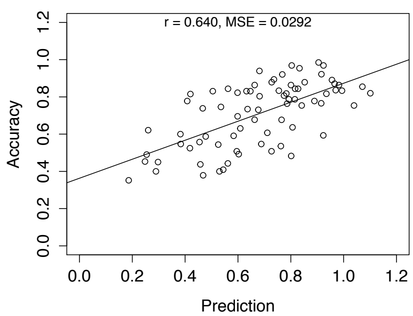

Figure 5: Leave-one-subject-out cross validation (LOOCV) performance prediction of 75 subjects compared with actual subject performance, using the weights of multiple regression analysis obtained from the rest of the subjects. A total of 4 predictors are 4 brain consistencies measured across selected brain voxels among the first 10 correctly solved problems for each of the 4 states. (r = .6404, MSE = .0292)

Conclusion

This study has shown that success in extending a human problem-solving procedure from familiar to unfamiliar problems is reflected in how consistent subjects' brain responses are. This consistency refers to both how consistently subjects respond to different problems (within-subject consistency), and how brain responses of one subject deviate away from the global average (between-subjects consistency), with the former one correlating more strongly with the subject overall performance. Within-subject brain consistency is most highly correlated with the subject performance when examining the early stage of the problem solving, which can serve as an effective neural predictor. During the later stage of the problem solving, we think that subjects start developing problem-specific strategies that decrease the brain consistency over time. We suggest that the relationship between subject performance and within-subject brain consistency is due to the perceived similarity between the familiar problems and the new problems.

Though previous studies have explored the relationship between consistency and subject performance during transfer of learning, our study is the first to identify such effect at the level of neural activity and use this consistency to predict performance. We also show that the accuracy of predicting subject performance has been further improved by selecting the most involved brain voxels rather than only using the predefined 12 regions. Both approaches lead to the same conclusion that subjects who have more consistent brain activation perform better.

References

Aly, M., & Turk-Browne, N. B. (2015). Attention stabilizes representations in the human hippocampus. *Cerebral Cortex*.

Anderson, J. R., Betts, S., Ferris, J. L., & Fincham, J. M. (2010). Neural imaging to track mental states while using an intelligent tutoring system. *Proceedings of the National Academy of Sciences*, *107*(15), 7018-7023.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*, 1036–1060.

Anderson, J. R., Bothell, D., Fincham, J. M., & Moon, J. (2014). The sequential structure of brain activation predicts skill. *Journal of Cognitive Neuroscience*.

Anderson, J. R., & Fincham, J. M. (2014a). Discovering the sequential structure of thought. *Cognitive Science*, *38*(2), 322–352.

Anderson, J. R., & Fincham, J. M. (2014b). Extending problem-solving procedures through reflection. *Cognitive Psychology*, *74*(0), 1 - 34.

Bunge, S. A., Wendelken, C., Badre, D., & Wagner, A. D. (2005). Analogical reasoning and prefrontal cortex: Evidence for separable retrieval and integration mechanisms. *Cerebral Cortex*, *15*(3), 239-249.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*, 487–506.

Epelboim, J., & Suppes, P. (2001). A model of eye movements and visual working memory during problem solving in geometry. *Vision Research*, *41*(12), 1561 - 1574.

Friston, K. J. (2007). *Statistical parametric mapping : the analysis of funtional brain images*. Amsterdam; Boston: Elsevier/Academic Press.

Geake, J. G., & Hansen, P. C. (2005). Neural correlates of intelligence as revealed by fmri of fluid analogies. *NeuroImage*, *26*(2), 555 - 564.

Gibbons, H., & Stahl, J. (2007). Response-time corrected averaging of event-related potentials. *Clinical Neurophysiology*, *118*(1), 197 - 208.

Gick, M. L., & Holyoak, K. (1987). *The cognitive basis of knowledge transfer*. Academic Press.

Glover, G. H. (1999). Deconvolution of impulse response in event-related bold fmri1. *NeuroImage*, *9*(4), 416 - 429.

Green, A. E., Fugelsang, J. A., Kraemer, D. J., Shamosh, N. A., & Dunbar, K. N. (2006). Frontopolar cortex mediates abstract integration in analogy. *Brain Research*, *1096*(1), 125 - 137.

Ischebeck, A., Zamarian, L., Schocke, M., & Delazer, M. (2009). Flexible transfer of knowledge in mental arithmetic — an fmri study. *NeuroImage*, *44*(3), 1103 - 1112.

Leberman, S., McDonald, L., & Doyle, S. (2006). *The transfer of learning: participants perspectives of adult education and training*. Gower Publishing.

Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fmri data. *Trends in Cognitive Sciences*, *10*(9), 424 - 430.

Obermaier, B., Guger, C., Neuper, C., & Pfurtscheller, G. (2001). Hidden markov models for online classification of single trial eeg data. *Pattern Recognition Letters*, *22*(12), 1299 - 1309.

Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. In *Proceedings of the ieee* (pp. 257–286).

Ross, B. H. (1984). Remindings and their effects in learning a cognitive skill. *Cognitive Psychology*, *16*(3), 371 - 416.

Thorndike, E., & Woodworth, R. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review*, *8*, 247-261.

Velazquez, J. L. P., & Wennberg, R. (2009). Denoising and averaging techniques for electrophysiological data. In *Coordinated activity in the brain* (Vol. 2, p. 165-189). Springer New York.

Wainer, H., & Messick, S. (1983). *Principles of modern psychological measurement: A festschrift for frederic m. lord*. Taylor-Francis.

Yu, S., & Kobayashi, H. (2006, May). Practical implementation of an efficient forward-backward algorithm for an explicit-duration hidden markov model. *Signal Processing, IEEE Transactions on*, *54*(5), 1947-1951.