Different Strategies in Solving Series Completion Inductive Reasoning Problems: An fMRI and Computational Study

Peipeng Liang, Xiuqin Jia (ppliang1979@gmail.com, xiuqin.jia@gmail.com)
Department of Radiology, Xuanwu Hospital, Capital Medical University, Beijing, China
Brain Key Lab of Magnetic Resonance Imaging and Brain Informatics, Beijing, China

Niels A. Taatgen (niels@ai.rug.nl)
Department of Artificial Intelligence, University of Groningen, Groningen, The Netherlands

Ning Zhong (zhong@maebashi-it.ac.jp)
Department of Life Science and Informatics, Maebashi Institute of Technology, Maebashi 371-0816, Japan

Kuncheng Li (lkc1955@gmail.com)
Department of Radiology, Xuanwu Hospital, Capital Medical University, Beijing, China
Brain Key Lab of Magnetic Resonance Imaging and Brain Informatics, Beijing, China

Abstract
Neural correlate of human inductive reasoning process is still unclear. Number series and letter series completion are two typical inductive reasoning tasks, and with common core component of rule induction. Previous studies have demonstrated that different strategies are adopted in number series and letter series completion tasks even the underlying rules are identical. In the present study, we examined cortical activation as a function of two different reasoning strategies for solving series completion tasks. The retrieval strategy, used in number series completion tasks, involves direct retrieving of arithmetic knowledge to get the relations between items. The procedural strategy, used in letter series completion tasks, requires counting a certain number of times to detect the relations linking two items. The two strategies require essentially the equivalent cognitive processes, but have different working memory demands (the procedural strategy incurs greater demands). The procedural strategy produced significant greater activity in areas involved in memory retrieval (dorsolateral prefrontal cortex, DLPPC) and mental representation/maintenance (posterior parietal cortex, PPC). An ACT-R model of the tasks successfully predicted behavioral performance and BOLD responses in DLPPC and PPC. The present findings support a general-purpose dual-process theory of inductive reasoning regarding the cognitive architecture.

Keywords: Number series; Letter series; Inductive reasoning; Adaptive control of thought-rational (ACT-R)

Introduction
Inductive reasoning, defined as inferring a general rule or relation from specific elements, is traditionally considered as one of the most important constitutes of human intelligence (Spearman, 1923). Several studies were performed to investigate the neural underpinnings of human inductive reasoning using different types of tasks, including sentiential (e.g., House cats have 32 teeth; Lions have 32 teeth; therefore, all felines have 32 teeth. Was the given conclusion plausible given the premises? Goel et al., 1997; 2004), figural (e.g., infer the rule underlying the figural stimuli consisting of novel animals. Goel et al., 2000) and numerical (e.g., number series completion tasks. Liang et al., 2007; Yang et al., 2009; Zhong et al., 2011; Jia et al., 2011) tasks. The recruitment of fronto-parietal regions together with their left lateralization are convergent reported in most of these studies, however, the detailed activation patterns are modulated by the heterogeneousness of the experimental task. Neural correlate of human inductive reasoning process is still unclear, and needs more experimental studies.

The series completion problem, including letter series (e.g., c e g ?) and number series (e.g., 3 5 7 ?), is a kind of typical inductive reasoning task (Pelligrino, 1985; Thurstone, 1938; Thorndike et al., 1986) and always used in general fluid intelligence (Gf) test (Cattel, 1963; Hayslip et al., 1995; Johnston et al., 2010; Redick et al., 2012). However, number series and letter series completion tasks are solved differently. Evidences demonstrated that different strategies (by definition, a strategy is a goal-directed procedure under the deliberate control of the participant (Rosenberg-Lee and Anderson, 2009)) were employed in number series and letter series completion tasks, in which each item in one has a same-rule counterpart in the other (Quereshi, 2001), and number series tasks are easier and more familiar than letter series tasks (Quereshi & Seitz, 1993; Quereshi & Smith, 1998). This has been confirmed by a recent pilot study with post-test oral report in our group, in which subjects are required to solve the two kinds of series completion tasks comprising items based on identical rules. A “retrieval” strategy is used in solving number series tasks, in which the relation between two adjacent items can be directly attained by retrieving the corresponding arithmetic fact from long-term memory (e.g., 11 13: 11+2=13, thus, the rule is +2). As to letter series tasks, a “procedural” strategy is adopted, in which participants require to step-wise count the corresponding adjacent letter of an item in order to find the relation linking two items (e.g., k n: k … l … m … n; thus, the rule is +3). The investigation of strategies using imaging has the potential to enrich our understanding the
neural substrates of inductive reasoning, in terms of locus, level, and duration of activity.

The retrieval and procedural strategy involve performing equivalent cognitive processes (retrieving of declarative memory to detect the relation between the adjacent two items, internal representation and maintenance, and response output), thus engage the same brain areas. However, the working memory demands differ between the two strategies. In the retrieval strategy (e.g., 11 13), participants directly get the relation (e.g., +2) between the adjacent two items by retrieving once the corresponding arithmetic fact (e.g., 11+2=13); While in the procedural strategy (e.g., k m), participants require to perform twice retrieval step by step (e.g., l is next to k, and m is next to l) and twice internal maintaining/updating a counter, and then the rule can be determined according to the counter. In this way, the procedural strategy incurs many more working memory demands than the retrieval strategy. Thus, the two strategies should differentially engage brain areas that are sensitive to working memory load. It is predicted that the two strategies can be differentiated by the extent of activity within the same brain areas, including the left prefrontal cortex, recruited in memory retrieval, and posterior parietal cortex, involved in mental representation (see a summarization in Anderson et al., 2008).

The goal of the current study was to employ computational cognitive modeling to make specific predictions about the strategy differences. Specifically, based on the aforementioned specification, we expect to demonstrate that the strategy difference can be distinguished neurally by differential engagement of the same brain areas. To make our predictions more precise (in terms of the timing and level of activity), we plan to build computational models of the experimental tasks using the adaptive control of thought-rational (ACT-R) cognitive architecture (Anderson, 2007). ACT-R proposes that cognitive process is the result of the independent activity of distinct modules which are coordinated by a central production system. The ACT-R models automatically generate predictions for activity in the ACT-R modules, which we could then compare with activity in the brain areas of interest. For the tasks in the present study, the differential engagement was mainly due to the differences in retrieval and maintenance demands between the two strategies. Two modules were thus of particular interest in this study: the retrieval module, which is responsible for the retrieval of declarative memories and linked to the lateral inferior prefrontal cortex (LIFPC) and the imaginal module, which is responsible for the encoding and maintenance of internal representation of the problems and linked to the posterior parietal cortex (PPC) (Anderson, 2007; Qin et al., 2004). As the procedural strategy incurs more demands of retrieval and maintenance, thus, we expected LIFPC and PPC have a greater response to the procedural strategy (used in letter series tasks) than the retrieval strategy (used in number series tasks). Together, we will test the two competitive theories empirically and computationally.

Materials and Methods

Subjects
Twenty-three paid healthy undergraduate and postgraduate students (11 females; 24.1 ± 3.7 years old) participated in the experiment. Written informed consent was obtained from each participant and this study was approved by the Ethics committee of Xuanwu Hospital, Capital Medical University.

Stimuli and experimental design
Four kinds of tasks were organized into a 2×2 factorial design (Table 1). The first factor was Content, consisting of two levels, number-related (24 inductions and 24 baselines) and letter-related (24 inductions and 24 baselines) tasks. The second factor was Task in which the first level was the induction condition consisting of series completion tasks (24 number series inductions and 24 letter series inductions) and the second level was the baseline condition (24 number judgment baselines and 24 letter judgment baselines). This yielded four types of tasks: number series induction (NumIR), letter series induction (LetIR), number judgment baseline (Is10) and letter judgment baseline (IsJ). In particular, interferential tasks, which are identical in pattern to inductions but without common rules (e.g., “1 3 8” or “a c f”), were included into NumIR and LetIR tasks based on a pilot study. There were twelve interferential tasks within 60 induction tasks in the current study, with six in NumIR tasks and another six in LetIR tasks.

For all kinds of tasks, there were three sequentially presented numbers or letters (e.g., “1 3 5” or “a c e”). Number-related tasks and letter-related tasks were matched for magnitudes and operations. All the letters involved were lowercase and within the range of a-z. Correspondingly, all the numbers involved were within the range of 1-26. Half of NumIR and LetIR tasks was forward (e.g., “1 3 5” or “a c e”; the rule is: +2) while another half was backward (e.g., “13 9 5” or “m i e”; the rule is: -4). In the answer options of NumIR and LetIR tasks, the distances between the correct and the false answer were less than 3. Half of Is10 and IsJ tasks were with the answer of “yes” while another half were with the answer of “no”.

<table>
<thead>
<tr>
<th>Table 1: Examples of experimental tasks.</th>
<th>Task</th>
<th>Options</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>NumIR</td>
<td>11 13 15</td>
<td>A. 18</td>
<td>B. 17</td>
</tr>
<tr>
<td>LetIR</td>
<td>t s r</td>
<td>A. q</td>
<td>B. p</td>
</tr>
<tr>
<td>Is10</td>
<td>14 23 10</td>
<td>A. yes</td>
<td>B. no</td>
</tr>
<tr>
<td>IsJ</td>
<td>n w j</td>
<td>A. no</td>
<td>B. yes</td>
</tr>
</tbody>
</table>

Stimuli Presentation
Stimuli from all conditions were organized into three sessions and presented randomly in an event related design. The order of sessions was counterbalanced among subjects. The beginning of a trial was signaled by a cue of the task type (“Finding a rule” for NumIR and LetIR tasks, “Is there
10” for Is10 tasks and “Is there ‘J’” for IsJ tasks) for 2 s. The numbers/letters then appeared on the screen with the first number/letter appearing at 2 s, the second at 2.6 s, and the last at 3.4 s. After the appearance of the third number/letter, subjects were instructed to press the left or right button (counterbalanced among subjects) when the answer was acquired. Three numbers/letters would remain on the screen for at most 6.6 s (since the presentation of the third number/letter), or the button-pressing response within 6.6 s would stop it. Subsequently, two answer options were displayed and subjects were asked to choose the correct answer by pressing the corresponding buttons (left button for “A”). Subjects were instructed to respond as accurately and quickly as possible and move to the next trial if the stimuli advanced before they could respond. The inter-stimulus interval (ISI) was 8 s and quick responses would leave more rest time within the trial. Thus, reaction times (RT) were recorded based on the first button response and accuracies were acquired based on the second button response.

MRI data acquisition
Scanning was performed on a 3.0 Tesla MRI system (Siemens Trio Tim; Siemens Medical System, Erlangen, Germany) with a 12-channel phased array head coil (see details in Liang et al., (2007)). This acquisition sequence generated 364 volumes for each session. The scanner was synchronized with the presentation of every trial.

Data Analysis

Data preprocessing
Data were analyzed using SPM5 software (http://www.fil.ion.ucl.ac.uk) and the preprocessing steps are identical to Jia et al. (2011).

fMRI analysis
Only correct response trials were included in data analysis (wrong response trials and interferential trials were not included). The epoch of interest is the duration between the presentations of the first number/letter to the first button response. The BOLD signal was modeled using the canonical HRF with time derivative and the RT as duration trial by trial.

For confirmatory analysis, region of interest (ROI) analyses focused on two predefined functional regions in ACT-R, LIPFC (TAL., -43, 23, 24) and PPC (TAL., -23, -63, 40) (Anderson, 2007). The statistical results were based on the beta-values (of general linear model implemented in SPM5) averaged within the two ROIs.

For exploratory analysis, we are primarily interested in the conjunction analysis to explore activations common to both strategies [(NumIR-Is10) and (LetIR-IsJ)]. The Task by Content interaction comparisons ([[NumIR-Is10]-[LetIR-IsJ]] and [[LetIR-IsJ]-[NumIR-Is10]]) were also executed to reveal the areas specific to each strategy. The activation reported survived a voxel-level intensity threshold of p < 0.05 corrected for multiple comparisons.

Results

Behavioral performance
We carried out analyses of variance for two factors: Content (number vs. letter) and Task (induction vs. baseline) on both RT and accuracy. Behavioral scores indicated that subjects performed the task in the expected manner. The main effects of Content and Task were significant for RT and accuracy: Response to induction tasks was significantly longer (F(1,22) = 70.10, p < 0.001) and less accurate (F(1,22) = 61.56, p < 0.001) than that of baseline tasks; Response to letter tasks was significantly longer (F(1,22) = 68.38, p < 0.001) and less accurate (F(1,22) = 69.94, p < 0.001) than that of number tasks. The interaction effect between Task and Content was significant for RT (F(1,22) = 69.24, p < 0.001) and accuracy (F(1,22) = 66.75, p < 0.001): The difference between LetIR and IsJ (i.e., LetIR-IsJ) (RT: F(1,22) = 30.59, p < 0.001; accuracy: F(1,22) = 73.81, p < 0.001) than the difference between NumIR and Is10 (i.e., NumIR-Is10) (RT: F(1,22) = 6.39, p < 0.05; accuracy: F(1,22) = 2.89, p =0.10). This indicated that the significant interaction effect is primarily driven by the strategy difference, i.e., the difference between LetIR and NumIR.

Imaging results

Confirmatory analysis We performed repeated measures Task by Content for the two ROIs. The two ROIs showed the identical patterns. The main effect of Content (F(1,22) = 35.46, p < 0.001 for LIPFC and F(1,22) = 14.23, p < 0.001 for PPC), Task (F(1,22) = 67.10, p < 0.001 for LIPFC and F(1,22) = 35.45, p < 0.001 for PPC) and the Task by Content interaction effect (F(1,22) = 29.64, p < 0.001 for LIPFC and F(1,22) = 4.60, p < 0.001 for PPC) were all significant. The post-hoc multiple comparisons showed that the interaction effect was driven by the strategy difference, i.e., the difference between letter series reasoning tasks and number series reasoning tasks (F(1,22) = 46.58, p < 0.001 for LIPFC and F(1,22) = 13.51, p < 0.001 for PPC). There were no significant difference between the two kinds of baselines (i.e., IsJ versus Is10; F(1,22) = 3.98, p =0.06 for LIPFC and F(1,22) = 1.85, p =0.19 for PPC).

Exploratory analysis Conjunction analysis indicated that both strategies showed common activation in the left dorsolateral prefrontal cortex (DLPPC, BA 46, 9) and bilateral superior parietal lobe (SPL, BA 7) extending to left inferior parietal lobe (IPL, BA 40) (Figure 1). Regions specific to LetIR [(LetIR-IsJ)-(NumIR-Is10)] were found in the left middle frontal gyrus (BA 9, 46), bilateral superior frontal gyrus extending into anterior cingulate cortex/medial frontal gyrus (BA 6, 8, 9, 32), inferior frontal gyrus/insula (BA 13, 45, 47), and bilateral superior parietal cortex (BA 7). No significant activations were found specific to NumIR [(NumIR-Is10) - (LetIR-IsJ)].
Having now reviewed the result of the experiment, we come to the question whether we can understand them in the framework of the ACT-R theory. The model we constructed primarily depends on the visual module to perceive the stimuli, the manual module to respond, the retrieval module to retrieve a fact from memory, and the imaginal module to encode and update its stored representation. Figure 2 presents the sequences of activity in the four modules, for a “+2” NumIR problem or a “backward next” relation LetIR problem solved by each of the strategies. Specifically, the underlying productions driving these modules are not shown. As the fronto-parietal network plays an important role in inductive reasoning, the retrieval and imaginal module are of special interest in the present study, in the following, we would illustrate the prediction of BOLD responses in the two modules.

The fit of the predictions of the model to the RT and ACC data are presented in Figure 3. The parameters were estimated to fit the behavioral data: a factor that scaled the time to retrieve a declarative memory fact (0.3 sec) and the time to modify the contents of the imaginal module (i.e., 0.2 sec). This leads to a predicted effect size (605.4 ms for NumIR, 2116.17 ms for LetIR, 551.25 ms for baseline conditions) that is highly similar to the observed effect size (586.4 ms for NumIR, 2221.4 ms for LetIR, 530.4 ms for Is10, and 606.1 ms for IsJ). Only the deviation of data is presented because the model makes identical predictions. It can be seen that the model produced a reasonable fit to the behavioral data.

In order to generate BOLD predictions, we convolved the module activity with a gamma function. As is typical (e.g., Boyton et al., 1996; Glover, 1999), if the module is engaged, it will produce a BOLD response at time units later according to the function:

\[ B(t) = \int_0^t D(x)H(t-x)dx \]

Once the timings of buffers actions are all set, we can predict the BOLD functions by estimating the magnitude parameter \( m \), the exponent \( \alpha \), and the latency scale \( s \) for each brain region. The estimates of these parameters and measurement of the quality of the prediction are given in Table 2. Figure 4 displays percent change of BOLD response (relative to the baseline defined by the average BOLD response of the first two scans and the last two scans), along with the prediction of the ACT-R model to be presented.

### Table 2: ACT-R parameters and the BOLD predictions.

<table>
<thead>
<tr>
<th>Imaginal</th>
<th>Retrieval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponent ((\alpha))</td>
<td>3</td>
</tr>
<tr>
<td>Scale ((s))</td>
<td>2.2</td>
</tr>
<tr>
<td>Magnitude</td>
<td>5.6</td>
</tr>
<tr>
<td>Correlation ((r))</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Figure 2: A schematic representation of the ACT-R model’s solution to the number- and letter-series problem of “11 13 15” and “t s r”, solved by retrieval strategy and procedural strategy, respectively.
Discussion

The primary purpose of the current study was to investigate the neural correlates of the two cognitive strategies in series completion inductive reasoning tasks. Both confirmatory and exploratory analysis indicated that DLPFC (L>R) and PPC (bilaterally) were commonly recruited in the two kinds of tasks, but with different extent. Additionally, the exploratory analyses identified additional regions in dorsomedial prefrontal cortex (DMPFC) and ventrolateral prefrontal cortex (VLPFC) was active for the procedural strategy but not for the retrieval one. These results suggest that some aspects of the behavioral signatures of strategies may be recovered from imaging data. We constructed the computational model to simulate participants’ behavior. The results showed that there was reasonable fitness between the model prediction and the empirical data.

In the current study, the significant co-activation of the left DLPFC and bilateral PPC is identified to be strategy-independent. These results replicated a recent fMRI study of number series completion task in our group (Jia, et al., 2011), and were consistent with previous studies of sentential inductive reasoning task (Goel and Dolan 2004) and figural inductive reasoning tasks (Induction minus Perceptual baseline; Goeland Dolan, 2000). This implies that these two regions may be the key regions involved in inductive reasoning. Moreover, the left DLPFC is more specific to sentential induction, as contrast to deduction (Goel and Dolan, 2004). The activity of the DLPFC is right lateralized in Goel et al. (2000), but is left lateralized in the other studies (Jia, et al., 2011; Goel and Dolan, 2004) and the current study, which may be ascribed to the different kinds of experimental materials (figure versus number, sentences).

As to the functional role of DLPFC and PPC, different interpretations were proposed. In the domain of inductive reasoning, the left DLPFC was related to the use of world knowledge in the generation and evaluation of hypotheses (Goel and Dolan, 2004), rule identification and extrapolations (Jia, et al., 2011), and the fronto-parietal network was more specific to rule identification (Jia, et al., 2011). (Note: In Goel and Dolan (2000), the authors are interested in bilateral hippocampus, which is specifically associated with rule inference, and right lateral orbital prefrontal cortex, which is associated with the task by difficulty interaction. The former is interpreted in terms of semantic encoding of novel stimuli, and the latter in terms of hypothesis selection. While the DLPFC (R>L) and bilateral PPC, identified in Rule Induction minus Perceptual Baseline, is not explained of their functional roles.) It is not surprise of these different explanations, as different domain and context were situated in.

In this study, with the help of computational modeling, we made a more general account. The DLPFC and PPC regions identified in the exploratory analysis, are largely overlapped the regions involved in the ACT-R LIPFC and ACT-R PPC. Therefore, consistent with the explanation in ACT-R theory, it was inferred that the left DLPFC is associated with memory retrieval of semantic information/knowledge and PPC is associated with mental representation of problem states (Anderson, 2007; Danker & Anderson, 2007). In the current study, although the procedural strategy and the retrieval strategy commonly recruited the left DLPFC and bilateral PPC, the strategy variation can be distinguished by the extent of cortical activity within these regions. The procedural strategy incurs more working memory demands (both the retrieval and maintenance demands) than the retrieval strategy. The ACT-R model fitted well with the experimental behavioral and imaging data, which demonstrated the validity of this general account.

![Figure 3: Data (red solid lines) and model fits (black dashed lines) for NumIR, LetIR, Is10, and IsJ.](image)

![Figure 4: The BOLD signals obtained for the prefrontal and parietal regions for NumIR and LetIR condition and the prediction of the ACT-R model of the task.](image)
In summary, by coupling this empirical work with computational modeling, we have deepened our understanding of what constitute the strategy variation (the retrieval strategy versus the procedural strategy) and quantify their predictions. These results support the general-purpose dual-process theory of inductive reasoning.

Acknowledgments

This work was supported by the Natural Science Foundation of China (No. 61105118), Beijing Nova Program (No. Z12111000250000). ACT-R code can be accessed by logging in ccureasoninglab@gmail.com (userid: ccureasoninglab; password: peipengliang).

References