Invariance detection in the brain: Revealed in a stepwise category induction task

Xueli Cai, Fuhong Li, Jing Wang, Hong Li

School of Psychology, Southwest University, Chongqing 400715, China
Research Centre of Brain and Cognitive Science, Liaoning Normal University, Chongqing 116029, China

A critical sub-process of category learning is detecting the invariance between categorical members. To examine brain activation associated with invariance detection at different steps of category learning, a stepwise category induction task was used in the present study. Within each trial, three stimuli were displayed sequentially, and participants were asked to learn the target category corresponding to the invariance among stimuli. Results revealed that invariance detection activated the fronto-parietal network. However, the frontal and parietal cortices functioned differently throughout the different steps of invariance detection. The left middle frontal gyrus (BA 9) was highly activated in both steps of invariance detection, but the posterior parietal regions, especially the right superior parietal lobule (BA 7), were more active in the final step of invariance detection, reflecting increased attention to the completion of category learning and the preparation for a subsequent response. Furthermore, a psychophysiological interaction analysis (PPI) revealed increased connectivity between the left middle frontal gyrus and the bilateral parietal cortex during the final step of invariance detection. Overall, the present findings imply the necessary role of the fronto-parietal network in variance detection.

1. Introduction

Category learning, the ability to recognize category membership of sensory stimuli, is critical for interpreting the meaning of events and preparing adaptive responses (Swaminathan and Freedman, 2012). The neural mechanisms of category learning have been broadly explored using different tasks, including dot-pattern prototype learning, “cat-dog” categorization, and rule-based category learning (Freedman et al., 2001, 2002, 2003; Hammer et al., 2009, 2010; Jiang et al., 2007; Li et al., 2009; Meyers et al., 2008; Miller et al., 2003; Pan et al., 2008; Pan and Sakagami, 2012; Seger and Cincotta, 2006; Seger and Miller, 2010; Sloutsky, 2010; Smith, 2008). For example, in the morphing continuum of “cat-dog (A–B)” categorization tasks, humans and monkeys are trained to categorize morphed images from A and B into an A-like category or a B-like category. Findings have shown that the inferior temporal cortex (ITC) is more involved in the analysis of currently viewed shapes, while the prefrontal cortex (PFC) shows stronger category signals (Freedman et al., 2003; Jiang et al., 2007; Li et al., 2009). In rule-based category learning tasks, monkeys have been trained to apply either a “same” or “different” rule to novel pairs of pictures, and results suggest that...
PFC neurons reflect abstract rule-based categorical distinctions (Muhammad et al., 2006; Wallis et al., 2001; Wallis and Miller, 2003).

Most recently, some studies attempted to assess cortical responses to sub-processes involved in category learning. For example, Hammer et al. (2010) found that detecting between-category differences was associated with the dorsal striatum and hippocampus, while detecting within-category similarities and differences was restricted to high-level visual brain areas. Garcin et al. (2012) demonstrated that similarity detection involved the anterior ventrolateral PFC (VLPFC) bilaterally with a right–left asymmetry, while abstraction of categories activated the left dorsolateral PFC (DLPFC).

Although neural correlates of category learning have been explored through the aforementioned paradigms, the neural basis of invariance detection, the critical sub-process of category learning, remains unaddressed. Vigo (2013) suggested that detecting invariance patterns in categorical stimuli is a necessary precursor to concept formation.

The purpose of the present study was to examine brain activation associated with invariance detection by using a three-step category induction task (CIT), which was developed according to previous studies dealing with hypothesis testing or category induction (Bigman and Pratt, 2004; Bruner et al., 1956; Chen et al., 2007; Levine, 1975; Li et al., 2013). During this task, participants were sequentially presented three stimuli that belonged to the same category and were asked to learn the target category corresponding to the invariance among stimuli. There were three perceptual attributes for each stimulus, but only one kept invariance across the three stimuli during each trial, which had been predetermined by the experimenter. When the first stimulus (S1) was presented, participants identified and remembered three attributes of the stimulus, each of which might have been related to the target category. When the second stimulus (S2) was presented, participants needed to detect the invariance (two shared attributes) between S2 and S1 while filtering out the variant attribute. When the third stimulus (S3) was presented, participants needed to further detect the invariance between S3 and the preceding two stimuli while filtering out the new variance. However, during the baseline task (BT), perceptual dimensions did not change across the three stimuli. For example, all three letters might be black, uppercase, and diagonal. The perceptual encoding and comparison involved during the BT was the same as for the CIT, but processes inherent to category induction, especially the process of filtering out the variant or conflict attributes, were not required during the BT (Fig. 1).

Invariance detection is necessarily accompanied by the inhibition of variance information (Pan and Sakagami, 2012; Garcin et al., 2012). It has been found that filtering relevant information from irrelevant information activates the prefrontal and parietal cortices (McCabe et al., 2010; McNab and Klingberg, 2008; Gazzaley and Nobre, 2012). Accordingly, we expected the fronto-parietal network to be associated with invariance detection. Moreover, invariance detection occurred only during the presentation of the second (S2) and third stimuli (S3) (Fig. 1); thus, we assumed that the fronto-parietal network might be significantly activated during the last two steps (S2 and S3) when compared to the first step (S1).

2. Results

2.1. Behavioral results

Reaction time (RT) and accuracy data were recorded for each trial during the CIT and BT. Accuracy was defined as the percentage of correct responses out of the total number of
trials during each task. Accuracy rates for the CIT and BT were 99% and 98%, respectively, which indicated successful performance by all participants.

RTs were defined as the time between the onset of the stimuli and key pressing. The RTs for the three letter stimuli and probe stimuli were calculated separately. The RT data for the letter stimuli were excluded (1.27%, 1.08%, and 0.59% for the first, second, and third letters, respectively) if participants made no response within the time limit.

As shown in Fig. 2, a repeated-measures ANOVA with a 2 (Task: CIT vs. BT) × 3 (Phase: phase 1, phase 2, phase 3) factorial design showed significant main effects of task and phase \([F_{\text{task}}(1,16) = 45.87, P < 0.001; F_{\text{phase}}(2,32) = 10.41, P = 0.002]\). RTs during the CIT (1494 ms) were substantially longer than those during the BT (1089 ms). RTs during phase 2 (1369 ms) were substantially longer than those during phase 3 (1081 ms). There was also a significant interaction effect \([F_{\text{interaction}}(2,32) = 24.10, P < 0.001]\). Further analyses showed that RTs for CIT2 (the second phase of the CIT) were significantly longer than those for CIT1 (the first phase of the CIT) and CIT3 (the third phase of the CIT) \([t(16) = 3.10, P = 0.001; t(16) = 3.65, P = 0.002]\), respectively. RTs for BT2 (the second phase of the BT) were significantly shorter than those for BT1 (the first phase of the BT) \([t(16) = 3.02, P = 0.008]\) and longer than those for BT3 (the third phase of the BT) \([t(16) = 6.31, P < 0.001]\). There was no significant difference between CIT1 and BT1 \([t(16) = −1.04, P = 0.315]\). RTs for CIT2 were longer than those for BT2 \([t(16) = 5.14, P < 0.001]\). RTs for CIT3 were also significantly longer than those for BT3 \([t(16) = 7.50, P < 0.001]\).

Probe RT data were excluded (approximately 0.29% of the total trials) if participants made no response within the time limit. Mean RTs for the probes during the CIT (1039 ms) were substantially longer than those during the BT (800 ms) \([t(16) = 9.07, P < 0.001]\).

### 2.2. fMRI results

Brain regions associated with activation for the main effect of task are displayed in Table 1 and Fig. 3. A main effect of task was found in the left middle frontal gyrus (BA 9) and left cingulate gyrus (BA 32). Further contrasts indicated that the left middle frontal gyrus (BA 9) was more active during the CIT than during the BT. Conversely, the left cingulate gyrus (BA 32) was more active in the BT than in the CIT.

A main effect of phase was found in the right postcentral gyrus (BA 85), right superior parietal lobule (SPL) (BA 7), left precuneus (BA 7), right middle frontal gyrus (BA 6), left precentral gyrus (BA 44), right anterior cingulate gyrus (BA 32), and right cingulate gyrus (BA 31). Further contrasts observed that the right postcentral gyrus (BA 5), right SPL (BA 7), left precuneus (BA 7) and right cingulate gyrus (BA 31) were more active in phase 3 than in phase 2, but there was no significant difference between phases 1 and 2.

A significant task × phase interaction was observed in a number of clusters in the following areas (all \(Fs > 11.48\); all \(Ps < 0.001\)): left middle frontal gyrus (BA 9), left medial frontal gyrus (BA 8), right SPL (BA 7), left inferior parietal lobule (IPL; BA 40), left SPL (BA 7), left postcentral gyrus (BA 2), left caudate, and right declive (Table 2). Further analyses of these areas indicated that there was no significant difference in brain activation between CIT1 and BT1. However, left middle frontal gyrus (BA 9), left IPL (BA 40), and left caudate were more active during CIT2 than during BT2 [all \(Ps < 0.001]\]. The left middle frontal gyrus (BA 9) and right SPL were more active during CIT3 than during BT3 [all \(Ps < 0.001\)].

![Fig. 2 – Reaction time data for the three phases (events) of the category induction task (CIT) and baseline task (BT). **P < 0.01. Error bars represent SE of the mean across all subjects.](image)

### Table 1 – Localization of activation during category induction.

<table>
<thead>
<tr>
<th>BA</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxels</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main effects of task</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left middle frontal gyrus</td>
<td>9</td>
<td>−45</td>
<td>6</td>
<td>36</td>
<td>1591</td>
</tr>
<tr>
<td>Left cingulate gyrus</td>
<td>32</td>
<td>−4</td>
<td>28</td>
<td>27</td>
<td>2600</td>
</tr>
<tr>
<td><strong>Main effects of phase</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right middle frontal gyrus</td>
<td>6</td>
<td>38</td>
<td>10</td>
<td>48</td>
<td>589</td>
</tr>
<tr>
<td>Left precentral gyrus</td>
<td>4</td>
<td>−29</td>
<td>−30</td>
<td>56</td>
<td>165</td>
</tr>
<tr>
<td>Right postcentral gyrus</td>
<td>5</td>
<td>32</td>
<td>−39</td>
<td>58</td>
<td>5301</td>
</tr>
<tr>
<td>Right superior parietal lobule</td>
<td>7</td>
<td>9</td>
<td>−66</td>
<td>58</td>
<td>1154</td>
</tr>
<tr>
<td>Left precuneus</td>
<td>7</td>
<td>−17</td>
<td>−73</td>
<td>56</td>
<td>841</td>
</tr>
<tr>
<td>Right anterior cingulate</td>
<td>32</td>
<td>2</td>
<td>41</td>
<td>12</td>
<td>1239</td>
</tr>
<tr>
<td>Right cingulate gyrus</td>
<td>31</td>
<td>2</td>
<td>−41</td>
<td>36</td>
<td>573</td>
</tr>
</tbody>
</table>

Note: BA = Brodmann’s area; CIT = category induction task; BT = baseline task; x, y, z = Talairach coordinates of the centroid of the region. False discovery rate-corrected, \(q < 0.05\).
during both the second and third phases, the BT elicited more activation in the left medial frontal gyrus (BA 8) than did the CIT [all Ps < 0.001]. In addition, among the three CIT phases, the left middle frontal gyrus (BA 9) was more active during CIT2 than during CIT1 [P < 0.001]. Conversely, more activity in the left medial frontal gyrus (BA 8) was found during CIT1 than during CIT2 [P < 0.001]. Stronger activation in the right SPL (BA 7) was found for CIT3 than for CIT2 [P < 0.001]. The left middle frontal gyrus (BA 9), left IPL, and right SPL were more active during BT3 than during BT2 [all Ps < 0.001] (Table 2 and Fig. 4).

In summarizing the imaging results for the ANOVA, especially the interaction effect, a network that includes both

---

**Table 2 - Brain regions showing an interaction effect of Task x Phase.**

<table>
<thead>
<tr>
<th>Area</th>
<th>BA</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxels</th>
<th>F(2,32)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Left middle frontal gyrus</td>
<td>9</td>
<td>-42</td>
<td>26</td>
<td>37</td>
<td>5905</td>
<td>23.46</td>
</tr>
<tr>
<td>Left medial frontal gyrus</td>
<td>8</td>
<td>0</td>
<td>38</td>
<td>37</td>
<td>13779</td>
<td>27.48</td>
</tr>
<tr>
<td>Right superior parietal lobule</td>
<td>7</td>
<td>30</td>
<td>-46</td>
<td>61</td>
<td>3108</td>
<td>24.65</td>
</tr>
<tr>
<td>Left inferior parietal lobule</td>
<td>7</td>
<td>-30</td>
<td>-52</td>
<td>43</td>
<td>370</td>
<td>17.19</td>
</tr>
<tr>
<td>Left superior parietal lobule</td>
<td>30</td>
<td>-36</td>
<td>-64</td>
<td>46</td>
<td>336</td>
<td>11.48</td>
</tr>
<tr>
<td>Left postcentral gyrus</td>
<td>2</td>
<td>-36</td>
<td>-37</td>
<td>58</td>
<td>693</td>
<td>11.73</td>
</tr>
<tr>
<td>Left caudate</td>
<td>-18</td>
<td>-10</td>
<td>28</td>
<td>295</td>
<td>13.99</td>
<td>15.33</td>
</tr>
<tr>
<td>Right declive</td>
<td>26</td>
<td>-68</td>
<td>-20</td>
<td>869</td>
<td>11.73</td>
<td>13.99</td>
</tr>
</tbody>
</table>

Note: BA = Brodmann’s area; x, y, z = Talairach coordinates of the centroid of the region. False discovery rate-corrected, q < 0.05.
frontal and parietal regions seems to play a significant role in category induction. More specifically, the left middle frontal gyrus (BA 9) might be closely related to cognitive processes involved in invariance detection. In addition, activation within posterior regions, mainly the right SPL, appeared to be more robust during the final phase of category induction. Unlike the left middle frontal gyrus, the left medial frontal gyrus (BA 8) was more active during the non-category induction task.

2.3. **Psychophysiological interaction analysis**

The PPI analysis, as shown in Fig. 5, revealed that activity in the left middle frontal gyrus (BA 9) from CIT1 to CIT2 was accompanied by a decreased functional interaction with the right postcentral gyrus (x=53, y=-12, z=48). However, activity in the left middle frontal gyrus (BA 9) from CIT2 to CIT3 was accompanied by an increased functional interaction with the right parietal lobe/supramarginal gyrus (x=59, y=-57, z=30) and the left IPL (x=-50, y=-59, z=42). At the chosen threshold criteria, there was no significant functional interaction from BT1 to BT2 and from BT2 to BT3.

3. **Discussion**

Humans can form a concept or learn a category by detecting invariance among categorical stimuli (Vigo, 2013). In order to investigate brain activation associated with invariance detection during category learning, we used a modified category induction task (Bigman and Pratt, 2004; Chen et al., 2007). Our behavioral data revealed longer RTs for the CIT than for the BT during the second and third phases of the task, reflecting the additional cognitive demands imposed during the CIT. During the CIT, participants had to detect the invariant attribute and inhibit the variant attribute in a step-by-step manner; finally, participants could get access to the target category after two information filtering steps. In contrast, there was no process of information filtering during the BT. The RTs during CIT2 were longer than those during CIT3, implying that filtering was more complex during CIT2 than during CIT3. Specifically, participants during CIT2 were required to filter two invariant attributes from one variant attribute while merely filtering one invariant attribute from one variant attribute during CIT3.

The fMRI results indicated that the CIT elicited greater brain activation in the fronto-parietal network. Compared to the BT, the CIT elicited more activation in the left middle frontal gyrus (BA 9), implying that this area may be related to the core sub-processes of category induction: detecting invariance and filtering out variance. The effects of phase and the interaction between task and phase indicated that the left middle frontal gyrus (BA 9) was more active during CIT2 and CIT3 events than during any other events, while the parietal regions were more active during CIT3 than during any other events. This suggests that different cognitive processes might be involved in different phases of category induction. In other words, the left middle frontal gyrus (BA 9) may be responsible for invariance detection and variance inhibition during category induction, while parietal activation may reflect increased attention to relevant representations when the process of category induction is completed. Although the main differences in brain activity between the CIT and BT in the present study seem to be restricted to the dorsolateral prefrontal and parietal cortices, both are part of the visuospatial working memory network (D’Esposito et al., 2000; Marvel and Desmond, 2010; Nystrom et al., 2000). The cognitive processes during the CIT are detecting invariance and filtering out variance, which differs from processes during a working memory task. In addition, memory load was the largest during CIT1, larger during CIT2, and smaller during CIT3. Consequently, fronto-parietal areas should be significantly activated during CIT1 if fronto-parietal activation is merely related to working memory. However, results of the present study do not
support this interpretation. That is, increased activation in the fronto-parietal network during CIT2 and CIT3 should not merely reflect differences in working memory involvement.

Several lines of evidence in monkeys and humans suggest that the lateral prefrontal cortex (LPFC), including the left middle frontal gyrus, is involved in encoding category information for a group of stimuli with perceptual or functional similarity during abstract categorization and category learning (Adams and Janata, 2002; Bunge et al., 2005; Devlin et al., 2002; Freedman et al., 2001, 2003; Garcin et al., 2012; Jiang et al., 2007; Li et al., 2013; Muhammad et al., 2006; Pan et al., 2008; Pilgrim et al., 2002; Reber et al., 2002; Shima et al., 2007; Tyler et al., 2001; Vogels et al., 2002; Wolfensteller and von Cramon, 2011). Some studies have found that the right LPFC contributes to the process of detecting physical similarities between items, whereas the left LPFC tends to be more involved in identifying conceptual relationships between items (Boroojerdi et al., 2001; Garcin et al., 2012; Green et al., 2006; Milton et al., 2009). For example, Garcin et al. (2012) found that similarity detection involves the anterior VLPFC bilaterally with a right–left asymmetry, while category abstraction activates the left DLPFC. Consistent with previous research, findings from the present study confirmed that invariance detection and variance inhibition were related to the left LPFC rather than the right LPFC. Unlike results from Garcin et al. (2012), the PFC activity in the present study was located in the DLPFC, which is close to the inferior frontal gyrus. Thus, it is possible that category induction in the

Fig. 5 – Results of the PPI analyses. (A) Left: reduced connectivity between the left middle frontal gyrus (BA 9) and right postcentral gyrus from CIT1 to CIT2. Right: mean activity in left middle frontal gyrus (Talairach coordinates (x, y, z): −42, 26, 37) is displayed as a function of mean activity in the right postcentral gyrus (Talairach coordinates (x, y, z): 53, −12, 48). (B) Left: increased connectivity between the left middle frontal gyrus (BA 9) and left inferior parietal lobule and right parietal lobule/supramarginal gyrus from CIT2 to CIT3. Right: mean activity in left middle frontal gyrus (Talairach coordinates (x, y, z): −42, 26, 37) is displayed as a function of mean activity in right parietal lobe/supramarginal gyrus (Talairach coordinates (x, y, z): 59, −57, 30) and left inferior parietal lobule (Talairach coordinates (x, y, z): −50, −59, 42).
present study was based on perceptual similarity rather than a more abstract biological category.

Although the CIT elicited activation in the fronto-parietal network, we found that frontal and parietal areas did not operate simultaneously during category induction. Frontal areas were activated throughout the whole process of category induction, while activation in the parietal cortex, especially the right SPL, increased gradually during the final step of category induction. To further assess the hypothesis that the left middle frontal gyrus interacts with posterior regions when participants made a response during the category induction task, a PPI analysis was performed to examine the functional integration of the left middle frontal gyrus and other regions during invariance detection. Both positive and negative interactions were observed (see Fig. 5). The left middle frontal gyrus had a decreased functional interaction with the right postcentral gyrus from CIT1 to CIT2 (the first invariance detection). It could be because functioning within the prefrontal cortex became increasingly more specific for extracting invariance and inhibiting variance during CIT2. This might have resulted in decreased collaboration with the posterior regions. From CIT2 to CIT3 (the second invariance detection), the left middle frontal gyrus had an increased functional interaction with bilateral parietal regions, and activity in the left middle frontal gyrus was stable during the two phases. It was possible that the left middle frontal gyrus showed a stronger modulation effect during CIT3 than during CIT2. During CIT3, one invariant piece needed to be detected, while one piece of variant information needed to be inhibited to complete category induction; hence, there was an increased functional interaction with the fronto-parietal network. Several studies have found that the fronto-parietal network is thoroughly involved in visual categorization and category-based decision making (Bigman and Pratt, 2004; Freedman and Assad, 2006, 2011; Seger and Miller, 2010; Swaminathan and Freedman, 2012; Vogels et al., 2002). Other studies have also indicated a close functional link between the PFC and the parietal lobe during figural inductive reasoning (Bunge, 2004; Bunge and Wallis, 2008; Christoff and Prabhakaran, 2001; Kroger et al., 2002; Silk et al., 2008). Stronger activation in bilateral posterior parietal lobes may be associated with the maintenance of previously created mental representations (Buschman and Miller, 2007; Vogels et al., 2002; Volle et al., 2010). In the present study, stronger activation in the posterior parietal regions might be related to increased attention to the completion of the CIT and preparation for a subsequent response. In short, the present results confirm the importance of functional connectivity between the PFC and posterior areas in promoting successful invariance detection during category induction.

It must also be noted that increased deactivation was observed in the left medial frontal gyrus (BA 8) during category induction. The task-induced deactivation found in this area may reflect default-mode network (DMN) activity (Buckner et al., 2008). The default-mode network is composed of a set of brain regions that become more active during low-demand tasks or rest (than during high-demand tasks) across a variety of situations (Buckner et al., 2008; Raichle et al., 2001; Raichle, 2010). This is thought to reflect certain types of cognitive processes that are more common during easy tasks or passive states and that may be suspended during performance of more effortful and goal-directed tasks. However, the nature of these processes is still a matter of debate and beyond the scope of the present study.

In conclusion, the present study adopted a modified CIT to reveal brain activation associated with critical sub-processes of category induction, specifically invariance detection. The fMRI results revealed that, compared to baseline task, invariance detection activated the fronto-parietal network. However, frontal and parietal regions had different roles during the different phases of category induction. The left mid-DLPFC became significantly activated during the two invariance detection steps. By contrast, the posterior parietal regions, mainly the right SPL (BA 7), were more active during the final step of invariance detection, reflecting increased attention allocated to the completion of category induction and the preparation of a subsequent response. A PPI analysis further revealed that activation in the left middle frontal gyrus was accompanied by an increase in functional integration with the bilateral parietal cortex during the final step of invariance detection. These findings imply that invariance detection is implemented by the fronto-parietal network, but future studies are needed to examine the effect of other factors (e.g., task complexity) on invariance detection.

4. Experimental procedures

4.1. Participants

Twenty right-handed, healthy volunteers took part in the experiment (10 men, 10 women; mean age, 22 years; range, 19–23 years). All participants met criteria for magnetic resonance imaging (MRI) scanning (i.e., no metallic implants, no claustrophobic, and a head size compatible with the custom head coil). In addition, participants had no known neurological or psychiatric injuries or disorders and were not taking any psychoactive medications or drugs. Data from three subjects were excluded before analysis because of unacceptable head motion or poor performance on the experimental tasks. All subjects gave informed consent, and the institutional review board of China’s Southwest University approved the study.

4.2. Materials and tasks

Stimuli labeled as edible biscuits were displayed sequentially in the center of a 17-inch screen. These biscuits were arranged to form fifteen letters (A, B, D, E, F, G, H, J, M, N, Q, R, T, U and Y). Stimuli varied along three perceptual dimensions, with two attributes for each dimension: color (white or black), letter case (uppercase or lowercase) and orientation (upright or diagonal). Variation in color, letter case, and orientation was randomized across letters and trials. Each letter was set in bold, and the size of the figures was approximately 6.65 cm in height and 4.84 cm in width. The color of the background was light gray (50% gray). The horizontal and vertical angles were both less than 3°.
4.2.1. Category induction task (CIT)
During the CIT, participants were asked to learn the target category (i.e., invariance of the letter biscuits). During each trial, three stimuli were presented one by one, and participants were informed that all presented biscuits were edible. For example, if a participant first saw a letter A, she/he may identify three attributes related to three possible hypotheses (e.g., black, uppercase, or diagonal biscuits were edible). A second letter B excluded one possible hypothesis (uppercase), leaving two other possibilities (black and diagonal) by comparing the second stimulus with the first. When a third letter, C, appeared, one more hypothesis (diagonal) was excluded, leaving “black” as the target. After the presentation of three letter biscuits, a probe stimulus, which consisted of a word written in Chinese, was presented. Participants were required to judge whether the written word indexed the target category (i.e., the only invariance among the three edible biscuits) by pressing one of three keys.

In order to examine patterns of brain activity during different steps of category induction, three phases were defined. Phase 1, or CIT1, corresponded to the presentation of the first letter (S1). During this event, three attributes related to three possible hypotheses could be registered and were to be memorized for subsequent perceptual comparison and invariance detection. Phase 2, or CIT2, corresponded to the presentation of the second letter (S2). During CIT2, comparing the differences between the current letter and the preceding one while detecting two invariant attributes would inhibit one variant attribute. Phase 3, or CIT3, corresponded to the presentation of the third letter biscuit (S3). During CIT3, one more variant attribute could be found, and one more hypothesis would be ruled out, leaving only one possibility (one invariant attribute) as the final answer.

4.2.2. Baseline task (BT)
The BT was the same as the CIT, with the exception that the perceptual dimensions did not change across three stimuli. For example, all three letters might be black, uppercase, and diagonal. When the second or third letter was presented, participants could not find the variant attributes between stimuli and could not eliminate any possible hypothesis. The perceptual encoding and comparison involved during the BT was the same as the CIT, but processes inherent to category induction, especially the process of filtering out the variant or conflict attributes, were not required during the BT.

4.3. Procedure
At the beginning of each trial, a fixation point (a black cross) appeared on the computer monitor for 500 ms. After a blank screen for 2–6 s, three letter biscuits were presented sequentially for a maximum of 4 s each or until the participants responded (by pressing key “4”). Between the presentations of letter biscuits, a blank screen was shown for 2–6 s. Finally, a written word (probe) was presented for a maximum of 4 s, or until participants responded. Participants had to judge whether the written words depicted the target category (i.e., the invariance of the edible biscuits). Participants were asked to press key “1” if the probe depicted the target category, “2” if the probe did not depict the target category, and “3” if they were unsure. The experimental procedures for the CIT and BT are shown in Fig. 6.

Participants completed four runs, each of which lasted eight minutes. The first two runs consisted of eight CIT trials and seven BT trials, and the other two runs consisted of seven CIT trials and eight BT trials, yielding a total of 30 trials per task for each subject. The two trial types (CIT and BT) were presented in a randomized order per run.

4.4. fMRI data acquisition
The fMRI data acquisition was performed using a Siemens TRIO 3.0 T full-body MRI scanner (Siemens, Erlangen, Germany). For each participant, anatomical images (256 × 256 × 176) with 1 mm × 1 mm × 1 mm resolution were obtained using a T1-weighted three-dimensional magnetization prepared rapid gradient echo (MPRAGE) sequence (inversion time, 900 ms; repetition time, 1900 ms; echo time, 2.52 ms; flip angle, 9°). Functional scanning used the echo planar imaging (EPI) flip angle (90°; field of view, 220 mm; in-plane resolution, 64 × 64; 0.99 mm gap; voxel size, 3.44 mm × 3.44 mm × 3 mm) with prospective acquisition correction (PACE), which helped to reduce head motion during data acquisition. Thirty-two axial slices were used to cover the whole cerebral cortex with no gaps, and slices were positioned along the anterior commissure–posterior commissure plane.

4.5. fMRI data analyses
Both pre-processing and statistical analysis of fMRI data were carried out using BrainVoyager QX (2.0). The first two volumes of functional scans were excluded from the analysis to ensure that steady-state tissue magnetization was reached. The remaining images were subjected to pre-processing, including three-dimensional motion correction, slice-time correction, spatial smoothing with a Gaussian kernel of FWHM (full width at half maximum) of 6 mm, and temporal high-pass filtering to remove nonlinear drifts of two or less cycles per time course. Then, functional EPI images from each participant were normalized by Talairach and Tournoux (1988) brain template in the dimension of 3 × 3 × 3 mm³. Finally, the resulting set of transformations was applied to the functional image volumes to form volume-time course representations, which would be used in subsequent analyses.

The blood oxygen level-dependent (BOLD) response of each event in the two tasks for each subject was estimated using a General Linear Model (Friston et al., 1995). BOLD responses (beta values) for the six events of each subject were submitted to a secondary group analysis following a random-effects model. Specifically, six predictors (corresponding to the three events in the two tasks) were used in the equation. In order to see whether reaction times across conditions explained some variance in the data, we defined an event-related parametric modulator based on standardized reaction time for each condition separately within each run (Buchel et al., 1998). Factorial design matrices were defined automatically from the created protocols. Statistical maps across all subjects were calculated for the main and interaction effects of the task and phase, and the results of these t-contrasts from each subject were then entered into a random-effects analysis at the group level. The threshold
for the random-effects maps was set at $q < 0.05$ to correct for false discovery rate (FDR) (Genovese and Wasserman, 2002).

4.6. Psychophysiological interaction (PPI) analysis

To assess the hypothesis that prefrontal activity involved in invariance detection interacts with posterior brain regions during category induction, we estimated the functional integration of invariance detection in a PPI analysis. PPI detects task-specific increases in the relationship between a seed region of interest and the rest of the brain, measured in terms of the strength of the regression of activity in one region on another (Friston, 2004; Friston et al., 1997). First, the deconvolved time course of activity was extracted in the left middle frontal gyrus from the task $/C_2$ phase factors identified as reflecting invariance detection and variance inhibition (a 6-mm radius sphere centered at the peak of activity in the group analysis; $x = -42$, $y = 26$, $z = 37$). We then calculated the product of this activation time course with the interaction terms of CIT2 versus CIT1, CIT3 versus CIT2, BT2 versus BT1, and BT3 versus BT2, to create the PPI. PPI analyses were carried out for each region of interest in each subject and then entered into a random effects group analysis (thresholded at $P < 0.001$ and a cluster size $> 10$ voxels).

Acknowledgments

This work was supported by the program for Liaoning Excellent Talents in University (LNET), and the National Natural Science Foundation (NNSF) Grants 31300844, 30370488, 31070985, 31271088, and 20100182110021.

References


