Chinese Science Bulletin 2003 Vol. 48 No.2 144—147

Neural correlates of "feeling-of-not-knowing": Evidence from functional MRI

LUO Jing[1], Niki Kazuhisa[2] & LUO Yuejia[1,3]

1. Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China;
2. Neuroscience Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Japan;
3. Lab Visual Information Processing, Chinese Academy of Sciences, Beijing 100101, China

Correspondence should be addressed to Luo Yuejia (e-mail: luoyj@psych.ac.cn)

Abstract **The neural correlates of "feeling-of-not-knowing" ("FOnK", i.e. the feeling-of-knowing judgments that accurately predicted "not knowing" or "misses" in the criterion test) were investigated by the event-related fMRI method through an RJR (recall-judgment-recognition) procedure that adopted unrelated word pairs as materials. Results revealed that, relative to the inaccurate "FOnK" predictions, the accurate ones were associated with activities in right ventral prefrontal cortex (PFC) and insula, the areas that were known to subserve "cue specification" in which the retrieval cues were converted into "descriptors" that could be used for direct memory search. This result implied that the accurate "FOnK" predictions relayed more on "cue specification" process than the inaccurate ones and was in consistent with the cue familiarity heuristic hypothesis of feeling-of-knowing.**

Keywords: event-related fMRI, feeling-of-knowing, cue specification.

Studies on feeling-of-knowing started from the 1960s[1]. In a typical feeling-of-knowing experiment, subjects were asked to provide an answer to a general knowledge question (e.g. the capital of a country) or to recall a target that had been paired with the cue in the learning phase. If subjects failed to provide the correct answer, they were asked to judge if they had a "feeling of knowing" on that solicited target, subjects were required to estimate the possibility of recognizing the correct answer among several selections, given recognition is usually easier than cued-recall. Finally, subjects did a criterion recognition task to testify their feeling-of-knowing predictions. There was a significant, but not very high, positive correlation between subjects' feeling-of-knowing predictions and their later recognition performance. This observation implied that people could still have a feeling of "I know that", even when they could not directly access the correct answer. Feeling-of-knowing is regarded as a kind of "memory on memories", i.e. the metamemory.

However, the neural basis of feeling-of-knowing is still unknown. Our knowledge on this topic mainly came from the studies on the brain-damaged person. For example, Shimamura and Squire[2] found that subjects with Korsakoff's syndrome were impaired on their feeling-of-knowing predictions. Since Korsakoff patients suffered from general cerebral atrophy, especially a frontal atrophy, it was reasonable to suppose that the feeling-of-knowing was based on frontal functions. But there was also evidence for brain areas other than frontal, temporal lobe, for example, underlying feeling-of-knowing judgments[3,4].

Neuroimaging techniques, especially event-related fMRI, provided a powerful way to investigate the neural correlates of feeling-of-knowing. Our recent neural imaging researches revealed that the accurate positive feeling-of-knowing predictions and the accurate negative feeling-of-knowing predictions might be mediated by distinct brain mechanisms[5]. The accurate positive feeling-of-knowing predictions ("feeling-of-knowing" or "FOK") were the metamemory judgments that accurately anticipated "knowing" (i.e. the "hits" in the final criterion recognition test), whereas the accurate negative feeling-of-knowing predictions ("feeling-of-not-knowing" or "FOnK") were the ones that accurately anticipated "not-knowing" (i.e. the "misses"). Previous feeling-of-knowing studies regarded the accurate "FOK" predictions and the accurate "FOnK" predictions as the same type of accurate metamemory predictions, but now, the brain imaging research revealed that the former was associated with robust activity in PFC, while the latter was accompanied with little. This implied that accurate "FOK" prediction was realized through an effortful retrieval process, whereas the accurate "FOnK" prediction was based on a "null" retrieval process, that was, little information was retrieved in the accurate "FOnK" prediction and subjects made the judgment of "I don't know" because nothing was accessed to the specific cue.

Researches on "FOnK" can provide new perspectives not only for feeling-of-knowing studies, but also for the understanding to the intelligence systems. Principally, an artificial intelligence system can know that it knows, but cannot know that it does not know. However, the studies on "FOnK" showed that there are at least two aspects in the human being's metamemory: on the one hand, he knows he knows; on the other, he also knows he does not know.

To investigate the mechanism of "FOnK", we contrasted the neural correlates of accurate "FOnK" predictions (i.e. participants said "no" to a specific item in the metamemory judgment, and also could not correctly recognize that item later) with that of inaccurate "FOnK" predictions (i.e. participants said "no" in the metamemory judgment, but correctly recognized that item later). Specifically, recent neuroimaging studies proved that the left PFC subserved the effortful and systematic memory retrieval, whereas the right side subserved the heuristic retrieval that was based on perceptual novelty/familiarity[6]. The heuristic retrieval called for a "cue specification" process in which retrieval cues were converted into "descriptors" that could be used for direct memory search. It was already known that the "cue specification" process could be embodied as the activation in ventral PFC and insula[7]. Based on this, we predicted that, relative to the inaccurate "FOnK" predictions, the accurate "FOnK" predictions would be associated with a more efficient "cue specification" process and lead to more activities in ventral PFC and insula.

1 Method

(i) Participants. Nine healthy, right-handed volunteers (20—22 years old, five females) recruited from the undergraduates participated in this experiment. They were interviewed several days before they attended the fMRI experiment and were given informed consent that followed the MRI ethics committee.

(ii) Materials. 160 unrelated word pairs that were composed of two-character, low frequency Japanese Kanji words were used as materials (for example, in the word pair "item-minister", "item" was the cue, the "minister" was the target). 80 of the word pairs were used as learning and memory items, the other 80 as lures in recognition.

(iii) Procedure. The experiment procedure followed the RJR paradigm in the feeling-of-knowing studies. There were three phases in the whole session: the learning phase, the cued-recall and FOK phase, and the recognition phase. Imaging was carried out in the cued-recall and FOK phase.

(1) Learning phase. Subjects were instructed to memorize each word pair that was presented to them so that they could recall the target word when was shown with the cue word. 80 word pairs were learned at a pace of 2.5 s per pair (2 s for item presentation, 0.5 s for unfilled delay) and subjects learned the list twice in a randomized order.

(2) Cued-recall and FOK judgments phase. 7 min after the end of the learning phase, the cued-recall and FOK phase started with fMRI scanning. The cue word was presented to the participants through a projector-screen-mirror system and they were asked to recall the target word that had been paired with it in the learning phase. The subjects were required not to speak or move heads during MR scanning; they pressed the keys of the response box that was attached on their right legs to indicate their judgments. The types of responses, but not the response time (RT) were recorded. There were three keys on the response box, if subjects successfully recalled the corresponding target word, they pressed the left key by the right index finger; if they could not recall the target, they were to press the middle key by the right middle finger when they felt they could recognize the correct answer from several candidates in the criterion recognition test, or, to press the right key by the right ring finger when they felt they could not recognize. Every cue word was presented for two seconds, followed by a 4.6-s unfilled delay. In the key-pressing condition, an asterisk, instead of a cue word, was presented at the same speed; subject was asked to press the three keys alternatively, that is, they pressed the left key to the first asterisk they saw, the middle key to the second, and the right key to the third. Four or three asterisk items presented successively and formed a key-pressing block. Ten cued-recall items presented successively and formed a cued-recall and FOK block. There were 16 blocks in all. 8 of them belonged to the cued-recall and FOK condition, 8 of them to the key-pressing condition.

(3) Recognition phase. 7 min after the end of the cued-recall and FOK phase, subjects performed a recognition test, in which 160 word pairs (half old and half new) were presented in a randomized order, subjects were asked to made an old/new judgment to each of these items in no more than 5 s.

(4) Sorting of the items. Based on the types of responses in the cued-recall and FOK phase ("successful cued-recall", "unsuccessful cued-recall and positive feeling-of-knowing", or "unsuccessful cued-recall and negative feeling-of-knowing") and in the recognition phase ("hits" or "misses"), items in the cued-recall and FOK phase were categorized into 5 types: PP items (positive-FOK, positive/ "hit"-recognition); NN items (negative-FOK, negative/ "miss"-recognition); PN items (positive-FOK, negative-recognition); NP items (negative-FOK, positive-recognition); and SC items (successful cued-recall, positive-recognition) (see Table 1). NN items and NP items were negative feeling-of-knowing predictions ("FOnK") in which subjects made "not knowing" judgments to the retrieval cues. In consistent with the metamemory predictions, NN items were not correctly recognized in the criterion test, whereas NP items were. We focused on these two types of items in the present research.

Table 1 Sorting of 5 types of metamemory items and their abbreviations

		Cued-recall and feeling-of-knowing judgment		
		Successful cued-recall	Unsuccessful cured-recall	Unsuccessful cured-recall
			Positive feeling-of-knowing	Negative feeling-of-knowing
Recognition	Hits	SC	PP	NP
	Misses		PN	NN

(iv) fMRI scanning. All scanning was performed on a 3.0 T MRI Scanner (GE Signa) equipped with EPI capability. 18 axial slices (5.5 mm thick, interleaved) were prescribed to cover the whole brain. A T2* weighted gradient echo EPI was employed. The imaging parameters were TR=3 s, TE=32 ms, FA=70 degrees, FOV=20×20 cm (64×64 mesh). To avoid head movement, we wore subject neck brace and inserted sponge pieces between the head and coil (also see our previous research[8]).

(v) Image analysis. Images were pre-processed (timeslice adjusted, realigned, normalized and smoothed) by SPM99. Then, imaging data of 9 subjects were estimated by a fixed effect model, using the Event Related Analysis of SPM99. Six types of events/items (PP, NN, NP, PN, SC and KP) were defined. The threshold was set at $P<0.0013$. The SPM coordinates for the standard brain from Montreal Neurological Institute (MNI) were converted to Talairach coordinates by a non-linear transform method (Image Homepage: www.mrc-cbu.cam.uk/Imaging/mnispace. html).

2 Results

We will focus our discussion on the neural correlates of accurate "FOnK" (NN) and inaccurate "FOnK" prediction (NP).

(i) Behavioral results. Gamma correlation between feeling-of-knowing judgments and recognition is 0.26 (95% CI Upper＝0.496，95% CI Lower＝0.026). This implied that the metamemory could predict memory performance in an above-chance level. Given "FOK" and "FOnK" were mediated by distinct neural basis, we considered them separately. 77% of the "FOK" judgments (including PP and PN items) accurately predicted the performance in criterion test, whereas only 47% of the "FOnK" items (include NN and NP items) did. In other words, just like the "FOK" judgments, in which most of the items could be correctly recognized in criterion test, most (53%) of "FOnK" items could also be correctly recognized. For this reason, the general predictive accuracy of feeling-of-knowing on recognition, as it was embodied as above-chance Gamma correlation, was mainly caused by the positive feeling-of-knowing judgments (i.e. the "FOK"), rather than the negative ones (i.e. the "FOnK").

(ii) Image results. The numbers of trials in NN and NP were almost the same (for NN, max. = 25, min. = 11, mean = 19.17, SD = 5.07; for NP, max. = 29, min. = 13, mean = 20.83, SD = 5.35), which enabled a balanced comparison between them. Relative to inaccurate "FOnK" predictions (NP items), the accurate "FOnK" predictions were associated with right ventral PFC activation that was peaked in insula (Table 2 and Fig. 1). Further examining on this activation proved two points. First, the peak point of signal change, which exhibited the maximum difference, was positive (Fig. 1); second, both NN and NP were associated with activation in this area, if we contrasted them with the baseline event (KP). This implied that this area participated not only in accurate "FOnK", but also in inaccurate "FOnK", but it was more active in the accurate ones. The reverse contrast, "NP items minus NN items", highlighted activities in left postcentral gyrus (BA 2), left cingulate gyrus (BA 31), and left precuneus (BA 31) (Table 2).

Table 2 Activation list highlighted in the contrasts of "NN minus NP" and "NP minus NN"

Contrasts	Talairach coordinates			*T*-value	Area
	x	*y*	*z*		
NN–NP					
	32	10	–2	3.08	right insula
NP–NN					
	–53	–20	30	2.86	left postcentral gyrus, BA 2
	–12	–23	40	2.67	left cingulate cortex, BA 31
	–22	–65	24	2.64	left precuneus, BA 31

$P<0.0013$, KE$>$10 voxels.

3 Discussion

Fletcher et al.[7] found that, relative to the retrieval that was based on the internal cues, the one based on the external cues led to more activities in right PFC and insula. So, this area was supposed to subserve the process of cue specification in retrieval. The present research observed the same activation in the contrast of "NN minus NP items", implying that the accurate "FOnK" predictions were associated with a more comprehensive cue specification process than the inaccurate ones. The process of cue specification has at least two cognitive functions. On the one hand, it enables the external cues imposed its control on the internal retrieval, to make the memory search to be a "field dependence"-like process, rather than a "field-independence"-like process. On the other hand, cue specification directed the intention of retrieval to the associations between the "cue" and the "target", rather than the "cue" itself, and for this reason, cue specification directed the endeavor of retrieval to focus on the search of "target". Relative to the inaccurate "FOnK" predictions, the accurate ones were accompanied by more cue specification processes, and so the memory search in accurate "FOnK" was better confined on the target that had been associated with the given cue. It was the sparseness of the retrieved target-related information that achieved accurate negative metamemory predictions.

The cue familiarity hypothesis of feeling-of-knowing proposed that this kind of metamemory judgments was made based on the information which one could get from the cue[9]. According to this hypothesis, cue specification could play a basic role in feeling-of-knowing judgments. The involvement of right PFC and insula in accurate "FOnK" predictions proved a cue specification process occurred in metamemory, and provided brain image evidence for the cue familiarity hypothesis of feeling-of-knowing. However, relative to the key pressing baseline, not only the accurate "FOnK" predictions but also the inaccurate ones were associated with activation in right PFC and insula, this implied that the cue specification

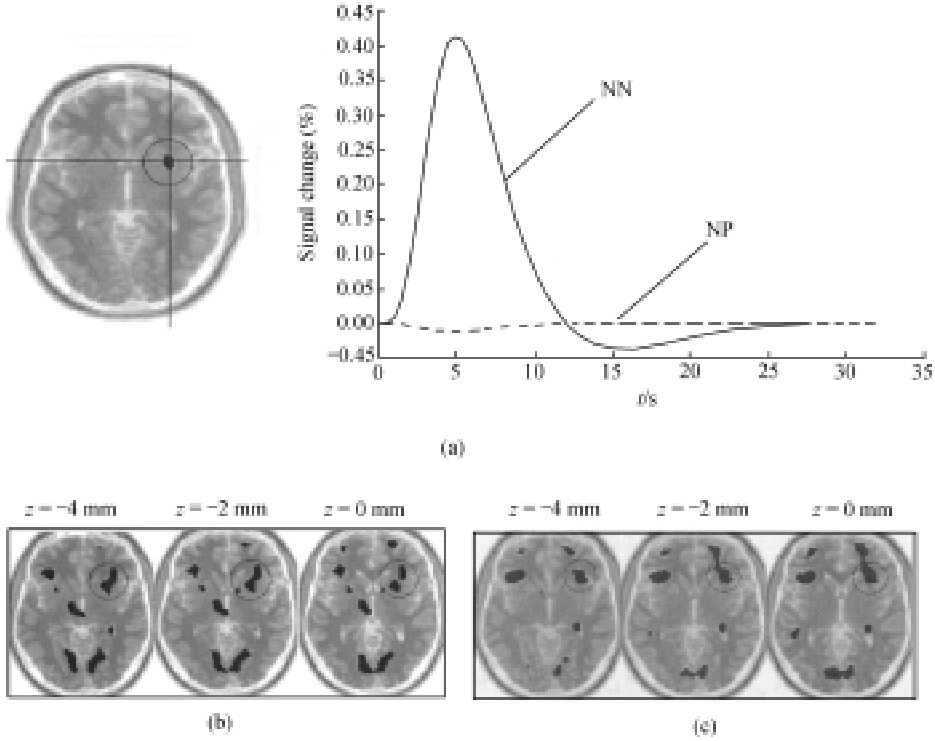

Fig. 1. Right PFC and insula activities that were revealed in "FOnK". White circles marked the activities located in right PFC and insula. (a) The contrast of "NN minus NP items"; (b) the contrast of "NN minus KP items"; and (c) "NP minus KP items". In (a), the cross in the left axial section marks the voxel that has the maximal value in the contrast (located in x, y, z = 32, 10, −2 of Talairach coordinates), the event-related plots on the right side are averaged signal change (%) of the best-fitting canonical hemodynamic response function (HRF) of 9 subjects from the maximal voxel. (b) and (c) show the axial sections of z = 0, −2, and −4 in the contrasts of "NN minus KP" and "NP minus KP" respectively. All of the figures show the averaged results of 9 subjects.

process was generally contained in "feeling-of-not-knowing" judgments, no matter they were accurate predictions or not. The accurate "FOnK" predictions contained a more comprehensive cue specification process than the inaccurate ones; it was this difference that decided the predictive accuracy of metamemory. We could not tell, however, what kind of difference lays between the accurate "FOnK" and the inaccurate "FOnK", although the comparison of these two had revealed significant difference in brain activity. Given that the cue specification process was generally involved in both the accurate "FOnK" and the inaccurate "FOnK" (relative to the baseline), and the present research was a neural correlation study that principally did not support any causal inference, we need further behavioral studies and the studies on brain-damaged person to discuss the possibility of dissociating the accurate "FOnK" and the inaccurate "FOnK".

Acknowledgments This work was supported by the National Natural Science Foundation of China (Grant No. 30270464), FAIPICS, the Multidisciplinary Research Program of the Chinese Academy of Sciences (Grant No. KJCX1-07) and by the Hundred Talents Program of the Chinese Academy of Sciences.

References

1. Hart, J. T., Memory and the feeling-of-knowing experience, J. Educ. Psychol., 1965, 56: 208.
2. Shimamura, A., Squire, L., Memory and metamemory: A study of the feeling-of-knowing phenomenon in amnesic patients, J. Exp. Psychol. Learn. Mem. Cogn., 1986, 12: 452.
3. Prevey, M. L., Delaney, R. C., Mattson, R. H., Metamemory in temporal lobe epilepsy: self-monitoring of memory functions, Brain Cogn., 1988, 7: 298.
4. Prevey, M. L., Delaney, R. C., Mattson, R. H. et al., Feeling-of-knowing in temporal lobe epilepsy: monitoring knowledge inaccessible to conscious recall, Cortex, 1991, 27: 81.
5. Luo, J., Niki, K., Imaging the metamemory: An event-related functional MRI research on feeling-of-knowing judgments (FKJs), Int. J. Psychol., 2000, 35: 40.
6. Nolde, S.F., Johnson, M.K., Raye, C. L., The role of prefrontal during test of episodic memory, Trends Cognit. Sci. 1998, 2: 399.
7. Fletcher, P. C., Shallice, T., Frith, C. D. et al., The functional roles of prefrontal cortex in episodic memory II, Retrieval. Brain, 1998, 121: 1249.
8. Niki, K., Luo, J., An fMRI study on the time-limited role of the medial temporal lobe in long-term autobiographic topographical memory, J. Cogn. Neurosci., 2002, 14: 500.
9. Metcalfe, J., Novelty monitoring, metacogntion, and control in a composite holographic associative recall model: Implications for Korsakoff amnesia, Psychol. Rev., 1993, 100: 3.

(Received October 14, 2002)