Research Report

Self-Reference During Explicit Memory Retrieval

An Event-Related Potential Analysis

Elena Magno[1] and Kevin Allan[2]

[1]*Institute of Neuroscience and School of Psychology, The Lloyd Institute, Trinity College Dublin, Dublin, Ireland, and* [2]*Cognitive Electrophysiology and Memory Laboratory, School of Psychology, College of Life Sciences and Medicine, University of Aberdeen, Aberdeen, Scotland, United Kingdom*

ABSTRACT—*Is there a specific neurocognitive system underlying the subjective sense of having a unitary continuous self across time? If so, it should be possible to isolate functions involved in the sense of self from those supporting mental activities that the self is currently engaged in. We report a study of real-time noninvasive recordings of the brain's electrical activity (event-related potentials, ERPs). We found a common neural signature that is associated with self-referential processing regardless of whether subjects are retrieving general knowledge (noetic awareness) or reexperiencing past episodes (autonoetic awareness). These ERP data are consistent with models of autobiographical memory that postulate a single locus of control over explicit memory for various kinds of self-related information. The temporal properties and scalp distribution of this novel self-reference ERP effect also suggest that it may be a neurophysiological correlate of self-related activation in medial prefrontal and parietal neocortical circuits identified in functional magnetic resonance imaging experiments.*

A central feature of human cognition is belief in one's unitary and continuous self across time. Although this belief is subjectively self-evident, its objective functional basis, if there is one, remains highly obscure (Gillihan & Farah, 2005). According to one influential framework (Tulving, 1983, 2001; Wheeler, Stuss, & Tulving, 1997), the belief is sustained by the ability to mentally time-travel in an *autonoetic* state of awareness that allows the qualities of specific past experiences to be recollected. In contrast to autonoetic states, *noetic* states provide access to self-knowledge abstracted from past experience (e.g., your belief that you are courageous).

Differences between noetic and autonoetic awareness have been studied in memory tasks by using the remember/know distinction. For each retrieval cue that subjects think they encountered during a previous study phase, they are asked to indicate whether they recollect some element of the episode in which the cue was presented (a *remember*, or R, judgment) or are merely aware that they had encountered the cue before (a *know*, or K, judgment; Gardiner, 2001; Tulving, 1985). These subjective judgments have been used with great success, notably to dissociate episodic from semantic memory (Tulving, 2001; Wheeler, Stuss, & Tulving, 1997), to refine theories of recognition memory (e.g., Yonelinas, 2002), to study the neuropsychology of amnesia (e.g., Aggleton et al., 2005; Levine et al., 1998), and, correspondingly, to explore the neural bases of memory function in vivo (e.g., Yonelinas, Otten, Shaw, & Rugg, 2005).

Although this research has illuminated memory function, Tulving and his colleagues have actually offered a tripartite framework for episodic memory (e.g., Tulving, 1983) that reaches beyond the encoding, storage, and retrieval of episodic content per se to include two further components: a cognitive self and the mental apparatus to represent it in subjective time during autonoetic states. Noetic and autonoetic states may thus differ in that the former give access to self-knowledge without the reference to a supporting episode from the subjective past that is part of the latter (e.g., Rosenbaum et al., 2005). In terms of Tulving's framework, noetic and autonoetic states may be clearly distinguished by the absence of mental time travel during noetic states (e.g., Wheeler et al., 1997). But from this tripartite perspective, autonoetic and noetic states may in fact have a common ingredient: self-related processing.

Address correspondence to Kevin Allan, Cognitive Electrophysiology and Memory Laboratory, School of Psychology, College of Life Sciences and Medicine, University of Aberdeen, Scotland, United Kingdom AB24 2UB, e-mail: k.allan@abdn.ac.uk.

Conway (2001; Conway & Pleydell-Pearce, 2000) has elaborated on Tulving's (1983) tripartite framework, proposing that the self may be conceived as a control function (the "working self") that establishes and maintains goals that regulate the ongoing formation of autobiographical memories (ABMs). Critically, each ABM may comprise representations of specific experience that can be recollected, as well as representations of self-knowledge abstracted from the personal past. Hence, in this view, all attempts to access self-relevant information, regardless of their noetic or autonoetic character, proceed through a common retrieval pathway involving the working self.

In the study reported here, we tested this conception of a functionally independent working self, combining electrophysiological neuroimaging with R/K judgments and a self-reference manipulation (e.g., Kelley et al., 2002; Macrae, Moran, Heatherton, Banfield, & Kelley, 2004). We used cue words (e.g., *spider*) to prompt an autonoetic subject group to recollect episodes that were either personal (e.g., "I saw a spider last night on television") or relevant to a friend (e.g., "Last night my friend said that she faints whenever she sees a spider") and to prompt a noetic subject group to generate facts that were relevant to either themselves or a friend (e.g., "I detest spiders" or "My friend cuddles spiders"). Subjects classified their responses using R/K judgments, allowing us to filter out responses of the autonoetic group that lacked recollection's subjective quality and responses of the noetic group that were "contaminated" by recollection.

Holding retrieval processes constant within each group, we attempted to isolate neural correlates of self-reference during autonoetic and noetic states, respectively. These correlates were revealed by analyzing scalp event-related potential (ERP) recordings of neural activity evoked by ABM cues, specifically, by comparing ERPs to cues that triggered retrieval of self-related information and cues that triggered retrieval of friend-related information. Although there is a considerable literature on ERP correlates of memory retrieval (see Rugg, in press), and on hemodynamic imaging of self-related processes (Gillihan & Farah, 2005), to our knowledge the fine temporal resolution of the ERP technique has not been exploited to disentangle the neural bases of self and memory, which may be concurrently activated. If access to self-relevant information, regardless of whether it produces noetic or autonoetic awareness, does involve a common functional component, then the same ERP self-reference effect would be expected in the two subject groups.

METHOD

Subjects

Subjects were 28 right-handed undergraduates at the University of Aberdeen, Scotland. They were screened for neurological, psychiatric, and psychological disorders and were not currently taking psychoactive medication. Fourteen subjects were randomly assigned to each group (noetic: mean age = 18.3 years, $SD = 0.7$; autonoetic: mean age = 19.4 years, $SD = 1.3$).

Materials and Procedure

The ABM stimuli were 160 visually presented low-frequency words (Kucera & Francis, 1967), divided into four 40-item lists. Each trial began with a 2,000-ms task cue, which indicated whether the subject should retrieve information relating to him- or herself or information relating to a specific friend. The task cue was followed by a 1,000-ms blank screen, then the ABM cue for 2,000 ms, and finally a 4,500-ms response period. In each list, 20 items were preceded by the cue "self," and 20 were preceded by the cue "friend"; these cues were presented in random order.

Subjects were seated in a sound-attenuated and electrically shielded recording chamber. Autonoetic subjects were instructed that if the task cue "self" appeared, they should attempt to retrieve and then describe out loud a specific personal episode relating to the following ABM cue. If retrieval was accompanied by a recollection, they were to press a joypad button labeled "R," and if no recollection was experienced, they were to press a different button, labeled "K." Examples of when "R" and "K" responses are appropriate were given, and feedback was also provided during a brief practice. Autonoetic subjects were also told that if the task cue "friend" appeared, they should use the ABM cue to recollect a memory concerning their friend, and give an appropriate "R" or "K" response.

Noetic subjects were instructed that for each ABM cue, they should generate a true statement of fact regarding themselves or a friend, depending on the task cue, as well as an accompanying "R" or "K" response. The groups were told that if no memory or suitable fact came to mind, then no response was necessary.

For scalp electroencephalogram recording and ERP formation, we followed the procedure outlined by Allan and Allen (2005). We used exploratory analyses of variance (ANOVAs) with factors of condition (self or friend) and recording site (29 different sites) to analyze mean ERP amplitude from successive 100-ms periods, beginning at 0 to 100 ms. We report ANOVAs for 20 sites (4 midline and 4 in each quadrant) during those latency intervals that showed consistent effects in these initial ANOVAs. Distribution analyses employed rescaled amplitude data (McCarthy & Wood, 1985), and all F ratios are Greenhouse-Geisser corrected for nonsphericity.

RESULTS

Behavioral Data

Table 1 shows that in the autonoetic group, the mean proportion of "R" responses exceeded the mean proportion of "K" responses, for both self- and friend-related memories, $t(13) = 6.31$, $p < .001$, $p_{rep} = .997$, Cohen's $d = 2.65$, and $t(13) = 4.88$, $p < .001$, $p_{rep} = .992$, $d = 1.74$, respectively. This is the

TABLE 1

Proportion of Trials on Which Autobiographical Cues Triggered Successful Memory Retrieval, by Group, Retrieved Content, and Recollective Experience

	Autonoetic group		Noetic group	
Content	R	K	R	K
Self-related	.61 (.18)	.19 (.13)	.08 (.07)	.77 (.18)
Friend-related	.51 (.18)	.24 (.12)	.08 (.07)	.76 (.22)

Note. Standard deviations are given in parentheses. Remember (R) judgments indicate retrieval was accompanied by recollective experience; know (K) judgments indicate it was not.

expected result, and this analysis acts as a kind of manipulation check. Also, self-related recollections were assigned an "R" judgment significantly more often than were friend-related recollections, $t(13) = 4.2$, $p = .001$, $p_{rep} = .985$, $d = 0.54$.

Almost the opposite pattern of results was obtained in the noetic group: Both self- and friend-related factual statements were assigned a significantly higher proportion of "K" than "R" responses, $t(13) = 12.24$, $p < .001$, $p_{rep} = .999$, $d = 5.07$, and $t(13) = 10.02$, $p < .001$, $p_{rep} = .999$, $d = 4.29$, respectively. This result, again, is as intended. However, the proportion of "K" responses did not differ between self- and friend-related statements, $t(13) < 1$, $d = 0.04$.

ERP Data

Neural correlates of self-reference were revealed by contrasts between ERPs evoked by ABM cues in the self versus friend conditions.[1] The autonoetic group's ERPs were formed exclusively on trials on which "R" responses were given, and the noetic group's ERPs were formed exclusively on trials on which "K" responses were given. Self-reference modulated neural activity, producing positive-going shifts in both groups' ERPs in the self condition, relative to the friend condition (see Fig. 1). These shifts emerged rapidly at midline sites in the autonoetic group, beginning at about 100 ms and lasting until the end of the epoch, at 1,944 ms. The noetic group's ERPs showed a similar but more temporally restricted self-reference effect around 800 to 1,200 ms. Figure 2 depicts the scalp distribution of the self-reference effects (i.e., the difference between self and friend ERPs).

Initial exploratory ANOVAs of the autonoetic group's data produced consistent main effects of self-reference throughout the 100- through 1,700-ms period. To compare this effect with that in the noetic group, we carried out focused analyses on data from the initial (100–400 ms), middle (800–1,200 ms), and final (1,200–1,700 ms) periods. For the autonoetic group, these analyses showed self-reference main effects at midline and quadrant sites during all three periods, consistently reflecting

[1]ERPs time-locked to the preceding task cues were insensitive to self-reference.

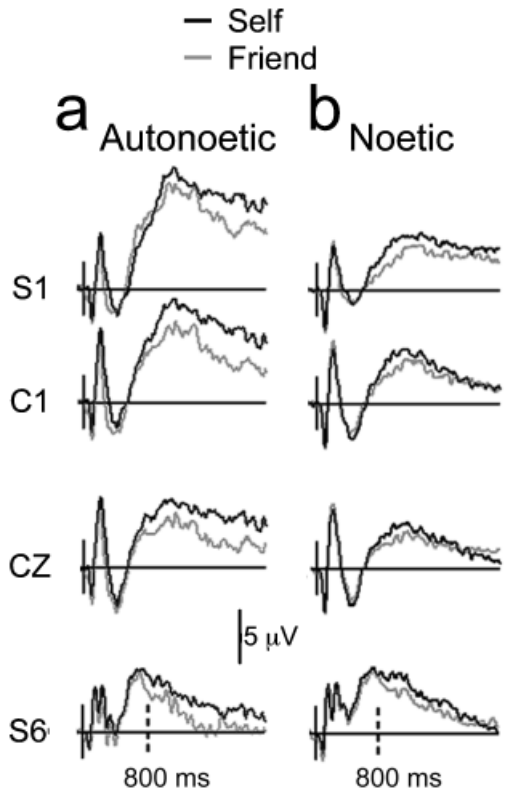

Fig. 1. Grand-average scalp midline event-related potentials (ERPs) time-locked to autobiographical memory cues that triggered successful retrieval in the (a) autonoetic and (b) noetic groups (the locations of the midline electrodes are indicated in Fig. 2e). Results are shown separately for self- and friend-related retrieved content. The ERPs for the autonoetic group were formed from a mean of 36.7 trials (range: 19–54) for self-related content and 30.5 trials (range: 16–52) for friend-related content; the ERPs for the noetic group were formed from a mean of 42.9 trials (range: 25–60) for self-related content and 42.7 trials (range: 23–66) for friend-related content. Positive is plotted upward.

the enhanced ERP positivity observed in the self condition, minimum $F(1, 13.0) = 4.75$, $p < .05$, $p_{rep} = .880$, $\varepsilon_p^2 = .268$. In the noetic group, only the 800- through 1,200-ms period yielded significant main effects of self-reference, at both midline and quadrant locations, again reflecting enhanced ERP positivity in the self condition, minimum $F(1, 13.0) = 5.03$, $p < .05$, $p_{rep} = .887$, $\varepsilon_p^2 = .278$. In a direct comparison, the magnitude of the self-reference effect did not differ significantly between the groups during the 800- through 1,200-ms period (i.e., the interval in which self-reference modulated both groups' ERPs).

A key remaining question, addressed by analyzing scalp distributions (Rugg & Coles, 1995), is whether the two groups' self-reference ERP effects are generated by the same or different underlying neural substrates. Visual inspection (Fig. 2) shows that throughout its time course, the self-reference ERP effect in the autonoetic group had a stable scalp distribution, focused just anterior and slightly to the right of the scalp vertex and extending posteriorly down the scalp midline. The self-reference effect in the noetic group, evident only during the 800- through 1,200-ms interval, was also diffusely distributed along

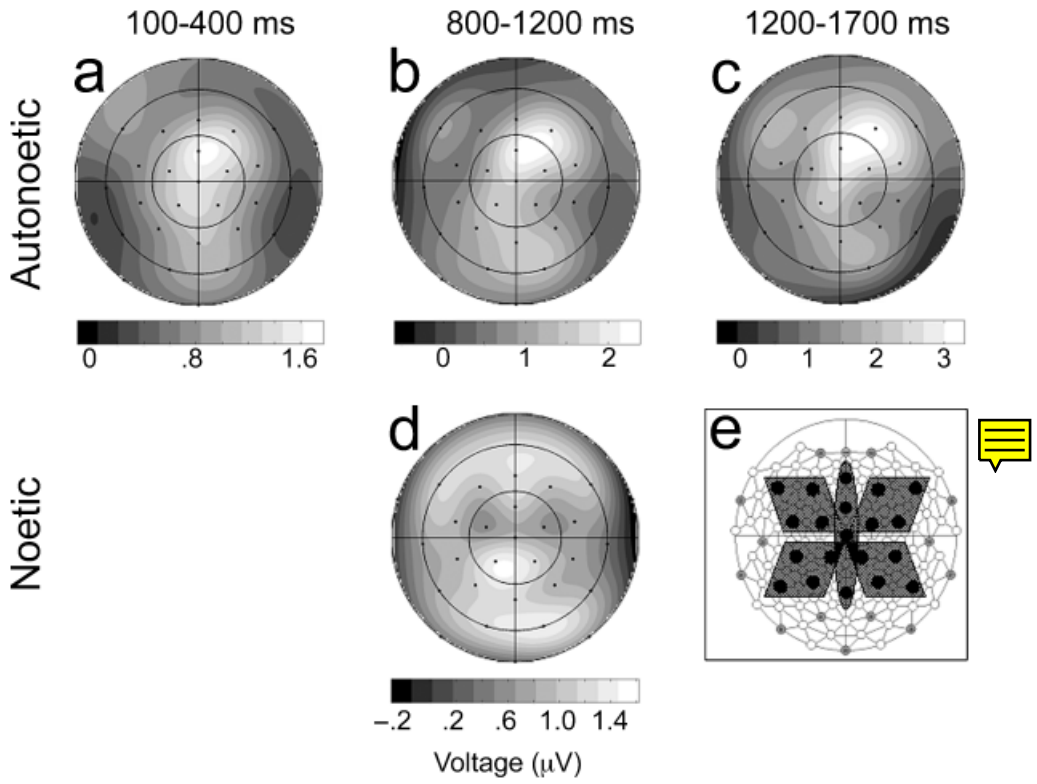

Fig. 2. Scalp distributions, by interpolated spline mapping, of the self-reference event-related potential (ERP) effects (i.e., ERP for self-related content minus ERP for friend-related content). The upper row (a, b, c) shows results for the autonoetic group in the three time intervals. The lower row shows (d) results for the noetic group in the middle time interval (i.e., the only interval in which the effect was significant), as well as (e) the locations of scalp electrodes for both the autonoetic group and the noetic group (the black shading highlights the midline and quadrant electrode sites selected for analyses).

the scalp midline, with some apparent differentiation between an anterior and a posterior component.

ANOVAs revealed no evidence of a significant difference in the distribution of the self-reference effect between the groups or across time within the autonoetic group; this was true both for analyses using data from all sites and for analyses using data specifically from the midline and quadrant subsets, maximum $F(1, 13.0) = 1.02$, $p = .32$. The absence of group effects indicates that the same configuration of neural generators produced the noetic and autonoetic self-reference ERP effects. In the autonoetic group, the absence of change over time confirms that these same generators were active consistently.[2]

DISCUSSION

The idea behind the present study was to shift focus away from mapping different memory functions during states of remembering and knowing, to obtain evidence relevant to a rather less clear issue—whether or not there is a cognitive self functionally independent of memory. Although there is long-standing, widespread, and growing interest in all aspects of self-related cognition (Gillihan & Farah, 2005), to our knowledge the present study is the first attempt to isolate and compare correlates of self-reference embedded in neural activity evoked during different retrieval states.

The attempt succeeded, insofar as we showed that a self-reference manipulation produced qualitatively identical ERP effects regardless of whether memory was accessed autonoetically or noetically. Hence, activity within a common neural substrate is modulated irrespective of whether access to self-relevant information triggers recollection or occurs without reference to a particular experience. The functional role of this newly identified component therefore transcends the particular retrieval "mode" (Tulving, 1983) in which information is explicitly (Schacter & Tulving, 1994) retrieved.

These data support Conway and his colleagues' proposal (e.g., Conway, 2001) that explicit access to self-relevant information involves a functionally independent executive-control component (termed the working self). In the present study, we may have isolated some functional component of the operation of this working self during long-term memory retrieval. A further indication that the present ERP self-reference effect is dissociable

[2]Because scalp ERPs do not homogeneously sample whole-brain activity, the absence of an interaction between self-reference and group or time does not preclude differences in the brain's functional state that other techniques (e.g., functional magnetic resonance imaging) may reveal.

from memory function per se is the fact that the various ERP components associated with long-term memory encoding and retrieval processes (Rugg, in press) do not resemble the correlate of self-reference identified here, although further work on this particular issue is required. The apparent scalp distribution and time course of the self-reference effect we obtained also differ from those of the P300 ERP modulations linked to recognition of one's own face and right-hemisphere function by Keenan and his colleagues (e.g., Keenan, Wheeler, Gallup, & Pascual-Leone, 2000), as well as from slow-wave correlates of autobiographical memory access identified by Conway, Pleydell-Pearce, Whitecross, and Sharpe (2003). These differences suggest that we have exposed a novel phenomenon.

The present neurophysiological data are perhaps most relevant to recent hemodynamic imaging (e.g., Fossati et al., 2004; Johnson et al., 2002; Kelley et al., 2002; Zysset, Huber, Ferstl, & von Cramon, 2002), which has consistently revealed that activity changes within regions of the medial prefrontal neocortex (mPFC), and also medial posterior cingulate (e.g., Lou et al., 2004; Seger, Stone, & Keenan, 2004), are triggered by self-reference manipulations. The distribution of the observed ERP self-reference effect, which spans anterior and posterior medial scalp, is consistent with such generator loci, although establishing the actual generator sites would require further empirical and modeling efforts. It is possible, however, to draw firm functional parallels between the present results and prior hemodynamic findings.

Notably, Levine et al. (2004) used functional magnetic resonance imaging to show that retrieval of episode-specific autobiographical information (e.g., what you ate for breakfast last Friday morning) and retrieval of information that generalizes across episodes (e.g., where you have breakfast every Friday) activate a common mPFC region, and that this activation is greater in magnitude when the information is episode-specific. These hemodynamic findings mirror the present neurophysiological results, which demonstrate that although similar self-reference effects occur during noetic and autonoetic states, these effects are faster acting and more temporally extended during autonoetic states.

The present findings therefore suggest that self-referential processing is required far earlier during autonoetic than during noetic retrieval, for reasons that are not clear. Intriguingly, the rapid onset of the autonoetic self-reference ERP effect has a parallel in reported transcranial magnetic stimulation (TMS) data. Lou et al. (2004) showed that TMS applied to the posterior medial parietal scalp selectively altered subjects' ability to determine whether personality-trait adjectives had previously been encoded in relation to the self. During this nominally autonoetic task, disruption occurred selectively when TMS was applied 160 ms after onset of the retrieval cue, suggesting a time-critical role for medial neocortical circuitry that coincides well with the initial onset of the autonoetic self-related ERP effect observed here.

Lou et al. (2004) speculated that interactions between anterior and posterior medial neocortical regions may underlie the unified subjective sense of self that we alluded to in our introduction. If there exists any objective underlying basis for this subjective construct within the brain's functional organization, then it should be possible to isolate functions involved in the sense of self from functions involved in the mental activities that the self engages in (Gillihan & Farah, 2005). Here we have provided evidence of a neural correlate that conforms exactly to this pattern: Specifically, we have shown that during two activities that involve accessing self-relevant information from long-term memory ("mental time travel" and accessing self-knowledge abstracted from past experiences), a common functional component may be engaged. Whether this component can itself be fractionated from other putative correlates of a putative unified self remains to be evaluated.

REFERENCES

Aggleton, J.P., Vann, S.D., Denby, C., Dix, S., Mayes, A.R., Roberts, N., & Yonelinas, A.P. (2005). Sparing of the familiarity component of recognition memory in a patient with hippocampal pathology. *Neuropsychologia*, *43*, 1810–1823.

Allan, K., & Allen, R. (2005). Retrieval attempts transiently interfere with concurrent encoding of episodic memories, but not vice versa. *Journal of Neuroscience*, *25*, 8122–8130.

Conway, M.A. (2001). Sensory-perceptual episodic memory and its context: Autobiographical memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *356*, 1375–1384.

Conway, M.A., & Pleydell-Pearce, C.W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, *107*, 261–288.

Conway, M.A., Pleydell-Pearce, C.W., Whitecross, S.E., & Sharpe, H. (2003). Neurophysiological correlates of memory for experienced and imagined events. *Neuropsychologia*, *41*, 334–340.

Fossati, P., Hevenor, S.J., Lepage, M., Graham, S.J., Grady, C., Keightley, M.L., et al. (2004). Distributed self in episodic memory: Neural correlates of successful retrieval of self-encoded positive and negative personality traits. *NeuroImage*, *22*, 1596–1604.

Gardiner, J.M. (2001). Episodic memory and autonoetic consciousness: A first-person approach. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *356*, 1351–1361.

Gillihan, S.J., & Farah, M.J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychological Bulletin*, *131*, 76–97.

Johnson, S.C., Baxter, L.C., Wilder, L.S., Pipe, J.G., Heiserman, J.E., & Prigatano, G.P. (2002). Neural correlates of self-reflection. *Brain*, *125*, 1808–1814.

Keenan, J.P., Wheeler, M.A., Gallup, G.G., & Pascual-Leone, A. (2000). Self-awareness and the right prefrontal cortex. *Trends in Cognitive Sciences*, *4*, 338–344.

Kelley, W.M., Macrae, C.N., Wyland, C.L., Caglar, S., Inati, S., & Heatherton, T.F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, *14*, 785–794.

Kucera, H., & Francis, W.N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Levine, B., Black, S.E., Cabeza, R., Sinden, M., McIntosh, A.R., Toth, J.P., et al. (1998). Episodic memory and the self in a case of isolated retrograde amnesia. *Brain*, *121*, 1951–1973.

Levine, B., Turner, G.R., Tisserand, D., Hevenor, S.J., Graham, S.J., & McIntosh, A.R. (2004). The functional neuroanatomy of episodic and semantic autobiographical remembering: A prospective functional MRI study. *Journal of Cognitive Neuroscience*, *16*, 1633–1646.

Lou, H.C., Luber, B., Crupain, M., Keenan, J.P., Nowak, M., Kjaer, T.W., et al. (2004). Parietal cortex and representation of the mental Self. *Proceedings of the National Academy of Sciences, USA*, *101*, 6827–6832.

Macrae, C.N., Moran, J.M., Heatherton, T.F., Banfield, J.F., & Kelley, W.M. (2004). Medial prefrontal activity predicts memory for self. *Cerebral Cortex*, *14*, 647–654.

McCarthy, G., & Wood, C.C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of variance models. *Electroencephalography and Clinical Neurophysiology*, *62*, 203–208.

Rosenbaum, R.S., Kohler, S., Schacter, D.L., Moscovitch, M., Westmacott, R., Black, S.E., et al. (2005). The case of K.C.: Contributions of a memory impaired person to memory theory. *Neuropsychologia*, *43*, 989–1021.

Rugg, M.D. (in press). Retrieval processing in human memory: Electrophysiological and fMRI evidence. In M.S. Gazzaniga (Ed.), *The cognitive neurosciences* (3rd ed.). Cambridge, MA: MIT Press.

Rugg, M.D., & Coles, M.G.H. (1995). The ERP and cognitive psychology: Conceptual issues. In M.D. Rugg & M.G.H. Coles (Eds.), *Electrophysiology of mind: Event-related potentials and cognition* (pp. 27–39). Oxford, England: Oxford University Press.

Schacter, D.L., & Tulving, E. (1994). What are the memory systems of 1994? In D.L. Schacter & E. Tulving (Eds.), *Memory systems 1994* (pp. 233–268). Cambridge, MA: MIT Press.

Seger, C.A., Stone, M., & Keenan, J.P. (2004). Cortical activations during judgements about the self and another person. *Neuropsychologia*, *42*, 1168–1177.

Tulving, E. (1983). *Elements of episodic memory*. Oxford, England: Oxford University Press.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist*, *25*, 1–12.

Tulving, E. (2001). Episodic memory and common sense: How far apart? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *356*, 1505–1515.

Wheeler, M.A., Stuss, D.T., & Tulving, E. (1997). Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness. *Psychological Bulletin*, *121*, 331–354.

Yonelinas, A.P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441–517.

Yonelinas, A.P., Otten, L.J., Shaw, K.N., & Rugg, M.D. (2005). Separating the brain regions involved in recollection and familiarity in recognition memory. *Journal of Neuroscience*, *25*, 3002–3008.

Zysset, S., Huber, O., Ferstl, E., & von Cramon, D.Y. (2002). The anterior frontomedian cortex and evaluative judgment: An fMRI study. *NeuroImage*, *15*, 983–991.

(Received 7/21/06; Revision accepted 10/3/06;
Final materials received 1/24/07)