HIPPOCAMPUS 16:491–494 (2006)

LETTER TO THE EDITORS

Aging, Memory, and Language in Amnesic H.M.

Donald G. MacKay

INTRODUCTION

To the Editors:

For many years, amnesic H.M. has been "a touchstone for research on amnesia and memory systems" (Manns, 2004), and the usual assumption is that H.M. exhibits a pure memory deficit, reflecting lesion-induced impairment that affects memory-encoding but not language systems. Related to this assumption, Kensinger, et al. (2001) made three significant claims that are examined here: (1) in 1999–2000 H.M. was unimpaired in processing preoperatively-acquired word-knowledge; (2) H.M.'s lexical information remained constant with aging from 1953 to 2000; (3) H.M.'s language processing was unimpaired at grammatical levels in 1999–2000. These Kensinger et al. claims are important because they strongly support systems-theory over competing theories in widespread use in the cognitive- and neurosciences. Under systems-theory, independent systems process language vs. memory. A language system comprehends verbal inputs, and then transmits the products of comprehension to a completely separate memory system for long term storage; a retrieval system later recovers the stored memory for transmission to a language production system, enabling verbal expression of the recovered memory. If correct, the claims of Kensinger et al. support these fundamental processing relations in systems-theories by dissociating the retrieval, comprehension, and production systems (undamaged in H.M.) from the memory storage system (damaged in H.M.).

However, all three claims of Kensinger et al. (2001) are in conflict with other evidence indicating that: (1) H.M. exhibits recent deficits in processing low-frequency (LF) words (James and MacKay, 2001; MacKay and James, 2001, 2002); (2) from 1983 to 1997, H.M. exhibited *exaggerated* age-linked declines in processing LF words that he knew and used correctly in 1970 (James and MacKay, 2001; MacKay and James, 2001, 2002); (3) H.M. exhibited syntax-level processing deficits in more than 30 sources of evidence from 1967–1973 (Lackner, 1974; MacKay, Burke, and Stewart, 1998a; MacKay, Stewart, and Burke, 1998b) to 1997–1999 (MacKay and James, 2001, 2002; MacKay, James, Taylor, and Marian, in press).

It is therefore important to examine procedural differences between the studies of Kensinger et al. vs. the studies of MacKay et al. that might explain their seemingly contradictory results. This letter does this for each claim of Kensinger et al. in turn and concludes that the two sets of results are compatible under a new theoretical framework known as binding theory (see e.g., MacKay and James, 2002; MacKay et al., in press).

Under binding theory, H.M. exhibits a *binding deficit* that predicts both *sparing* and *impairment* of a parallel nature in language and memory (see e.g., MacKay et al., in press, for detailed theoretical rationale). As applied to H.M.'s lexical knowledge in 2000, 47 years postlesion, H.M.'s binding deficit predicts selective impairment in processing LF but not familiar or high-frequency (HF) information. Second, for information encountered since his lesion, H.M.'s binding deficit predicts selective impairment for rapid (single trial) learning but not for slow learning, resulting from massive repetition over many trials or encounters with new lexical information. Third, binding theory predicts an age × lexical deficit interaction, with *exaggerated* age-linked declines from 1991 (age 65) to 2000 for processing LF but not HF lexical information. The data of MacKay–James supported all three of these binding-theory predictions.

Consider now the first claim of Kensinger et al. (2001)—H.M.'s lexical processing is unimpaired. To support this claim, Kensinger et al. presented three sources of data. Source one was experimental. However, the relevant experiments of Kensinger et al. mainly involved information that most certainly has been highly familiar to H.M. since childhood, e.g., recognizing and producing names of familiar objects, spelling familiar (mostly HF) words, identifying place names associated with highly familiar landmarks or events, e.g., Alamo-Texas, multiple-choice category recognition for familiar birds, fruits, furniture, and insects, and production of highly familiar plural and past-tense forms, e.g., boy–boys and dig–dug. These results of Kensinger et al. simply reillustrate H.M.'s preserved function for highly familiar lexical information and do not discriminate between systems theory vs. binding theory, because binding theory predicts *selective* deficits that do not include HF information.

By contrast, the research of MacKay–James satisfied this and other preconditions for discriminating between systems theory vs. binding theory. Consider the meaning–definition experiment of James and MacKay

Psychology Department, University of California, Los Angeles, California
Grant sponsor: Samuel A. MacKay Memorial Research Fund.
*Correspondence to: Donald G. MacKay, Psychology Department, University of California, Los Angeles, CA 90095-1563.
E-mail: mackay@ucla.edu
Accepted for publication 31 January 2006
DOI 10.1002/hipo.20176
Published online 29 March 2006 in Wiley InterScience (www.interscience.wiley.com).

TABLE 1.

Parallels Between Misderivation and Malaprop Errors in H.M.'s Definitions of LF Words and Pseudowords in the Work of James and MacKay (2001)

	Error type	
	Misderivation errors	Malaprop errors
Examples of H.M.'s definitions of LF words	*Lentil:* ''A combination word, in a way, from lent and till… (meaning) area and time of''	*Squander:* ''to take things as one's own, other persons' things''
Examples of H.M.'s definitions of pseudowords	*Unmelt:* ''something to stay stiff or not melted''	*Pediodical:* ''about the same thing as periodical''

(2001). In this task, H.M., cerebellar patients, and memory-normal controls matched with H.M. for 1997 age, IQ, and education, simply defined words that were either HF (e.g., What does *payment* mean?) or LF (e.g., What does *squander* mean?). Consistent with binding theory, systems theory, and H.M.'s nondeficits for highly familiar lexical information in the results of Kensinger et al., H.M. correctly defined 2% more of HF words than controls, a nonsignificant difference. However, as predicted under binding theory but not systems theory, H.M. correctly defined 46% fewer of LF words than controls, a 3.8 standard deviation (SD) deficit. Performance of the cerebellar patients ruled out H.M.'s cerebellar damage as the basis for his deficit in LF word comprehension. More importantly, longitudinal comparisons ruled out failure to learn as the basis for H.M.'s deficit: H.M. correctly used LF words in 1970 that he could not use correctly in 1997, an unlikely outcome given the failure to learn during childhood.

H.M. also made two types of definitional errors that clarified the nature of his deficit for LF information—misderivations and malaprops. Misderivations were definitions based on inappropriate morphological analyses of a stimulus. For example, H.M. defined the LF word *lentil* as a combination of (pseudocomponents) *lent* and *till* (see Table 1). Malaprops were definitions that matched a familiar, phonologically-similar word better than the target stimulus. For example, H.M.'s definition of *to squander* fits *to plunder* better than *to squander* (see Table 1). Misderivations and malaprops occurred for H.M. but not for controls, and this indicated that LF words had become meaningless to H.M. because he produced similar errors when asked to define truly meaningless pseudowords that he mistakenly claimed were words (see Table 1).

Consider now the second source of alleged support of Kensinger et al. (2001) for unimpaired lexical knowledge—comparisons of H.M.'s pre- vs. postoperative scores on the vocabulary, information, similarities, and interpretation subtests of the W-B I, their most frequently administered IQ test, and the only test administered both pre- and postoperatively. This evidence seems to contradict binding-theory claims regarding lesion-induced impairment of lexical knowledge. However, this surface appearance is deceptive. The reason is that H.M.'s preoperative IQ tests were administered shortly before his operation, when H.M. was experiencing high-frequency grand mal attacks and minor seizures up to 10 times a day, when he was on heavy and varied anticonvulsant medication, and when his mental state was "confused" according to Scoville (1968). Confirming the special status of H.M.'s preoperative scores, Scoville reported postoperative reductions in H.M.'s seizure-frequency and postoperative improvement in his intellectual function, including complete clearing of his confusion and "a slight but definite improvement in memory function." H.M.'s preoperative IQ scores must therefore be discarded when evaluating the possible effects of his lesion and when comparing H.M.'s longitudinal performance over time to evaluate the possibility of exaggerated age-linked decline. This evidence does not therefore contribute to the issue of lesion-linked lexical impairments and does not discriminate between systems theory vs. binding theory.

The third source of alleged support was age-linked. The analyses of Kensinger et al. (2001) indicate no reliable decline relative to age norms in H.M.'s performance from 1953 to 2000 on the vocabulary, information, similarities, and interpretation subtests. On the surface, this finding seems to contradict both binding theory and the extensive evidence indicating exaggerated age-linked declines in H.M.'s ability to process LF words (James and MacKay, 2001; MacKay and James, 2002). However, four procedural factors render this surface appearance deceptive. Factor one was just described—about 17% of the data in the analyses of Kensinger et al. consisted of unrepresentative *preoperative* scores.

Factor two was that H.M. has repeated virtually the same IQ tests at least 20 times since his 1953 lesion (Kensinger et al., 2001) so that aging and massive repetition are confounded in H.M.'s test scores but not in the age norms. Because H.M. benefits from massive repetition under binding theory, H.M.'s longitudinal subtest scores are therefore invalid for evaluating either the effects of aging or the impairment in 2000 relative to age norms.

Factor three was that Kensinger et al. (2001) based their age-linked conclusions on null results without sufficient statistical power to detect age-linked declines. For example, age- and test-linked variability was high (especially for the comprehension and vocabulary subtests, see Kensinger et al. 2001, Fig. 8, p. 356), and sample sizes were small (e.g., with $N = 1$ per test score, N was 6 for the most frequently administered test). In short, inadequate statistical power rendered the null results of Kensinger et al. meaningless, especially given a strong trend toward exaggerated age-linked decline on vocabulary subtests (discussed shortly).

Factor four concerned familiarity and age-of-acquisition. The information, similarities, and interpretation subtests depend virtually exclusively on retrieving familiar information that H.M. acquired before 1953 and has been using in everyday life since then, processes that remain constant with aging in binding theory and in empirical data for normal older adults (see e.g., MacKay and Burke, 1990) and H.M. (James and MacKay, 2001; MacKay and James, 2001, 2002). For example, the information subtest involves retrieval of semantic information normally acquired during childhood for questions such as "What is a thermometer?" "How many weeks are there in a year?" and "What does the heart do?" The similarities subtest involves retrieval of information normally highly familiar well before H.M.'s age at the time of his lesion, e.g., questions such as "How is a daily paper and a radio alike or similar?" (correct response: "sources of information"; less general responses such as "both convey news" do not get full credit); and "Why are eggs and seeds alike or similar?" (correct response: "young come from both"; More general responses such as "both are means of reproduction" do not get full credit). The interpretation or general comprehension subtest, despite its name, does not tap language comprehension per se but rather what might be called "appropriate sociocultural responses" involving familiar information normally acquired during childhood. A typical question is, "What would you do if you find a letter in the street that is sealed, addressed, and stamped?" (correct answer: "put it in the nearest mailbox"; other more imaginative or context-specific answers get no credit, e.g., "take it to the person to whom it is addressed," or partial credit, e.g., "take it to the police or post office"). Because binding theory does not predict age-linked deficits for highly familiar lexical information acquired before his lesion, H.M.'s performance on the information, similarities, and general comprehension subtests is irrelevant to binding theory.

This is not to say that H.M.'s test score data cannot in principle discriminate between systems theory vs. binding theory. Binding theory predicts impairment in H.M.'s postlesion performance for aspects of the W-B I that require novel encoding or creative analyses involving the manipulation of new information. H.M.'s scores on the arithmetic reasoning (AR) subtest (not reported in (Kensinger et al., 2001) represent a case in point. Binding theory predicts no significant deficits for the set of AR related word-problems that can be solved using information acquired during childhood, e.g., "How much is four dollars and five dollars?" but predicts major deficits for other problems that require formation of never previously encountered representations or novel application of basic mathematical propositions. Examples are: "If 8 men can finish a job in 6 days, how many men will be needed to finish it in half a day?" and "A man bought a second hand car for two-thirds of what a new would cost. He paid $400 for it. How much would a new car cost?". If H.M.'s age-adjusted scores fail to exhibit this predicted interaction between familiar vs. novel AR problems, this would call into question the binding-theory claim that H.M.'s lesion has impaired the nonrepetitive encoding and manipulation of new information.

Finally, the familiarity/repetition factor applies especially to vocabulary subtests, in which defining LF words is central to success, because the frequency-graded structure of these subtests makes deficits and age-linked declines especially difficult to detect in H.M. For example, participants taking the W-B I vocabulary subtest define words with progressively decreasing frequency of use, from extremely HF words, e.g., No. 1 "apple" (correct response "fruit"), and No. 2 "donkey" (correct response "animal"), to extremely LF words, e.g., No. 24 "belfry," No. 27 "pewter," No. 37 "amanuensis," No. 38 "proselyte," and No. 42 "traduce." Because it is unlikely that H.M. knew such extremely LF words at the time of his 1953 lesion (age 26), his performance for these words will be close to floor and invariant with aging under binding theory as H.M.'s binding deficit makes LF words encountered after his operation extremely difficult to learn. However, H.M.'s performance for HF words that he had learned as a child and has been using frequently since 1953 will be age-invariant under binding theory because aging does not impair frequently used information (see e.g., MacKay and Burke, 1990). In short, the selective nature of H.M.'s word-meaning deficits would make the effects of aging extremely difficult to detect via W-B vocabulary subtests even given adequate statistical power. Nonetheless, H.M.'s scores on vocabulary subtests administered from 1958–1998 exhibited a strong and marginally reliable decline ($P < 0.10$; Kensinger et al., 2001).

Four factors therefore explain the otherwise mysterious conflict between the meaning–definition results of James and MacKay (2001) vs. H.M.'s age-normed vocabulary scores—inclusion of H.M.'s unrepresentative 1953 vocabulary score in the analyses of Kensinger et al. (2001) worked against observing longitudinal declines relative to age norms; aging was confounded with effects of repetition in H.M.'s longitudinal IQ scores but not in the age norms; power for detecting age effects was inadequate; and the frequency-graded structure of W-B vocabulary subtests made age effects especially difficult to detect in H.M. Any one of these factors would render the data of Kensinger et al. invalid for evaluating effects of aging. Note however, that none of these problems apply to the results of James and MacKay (2001), in which H.M. exhibited no age-linked declines relative to age-matched controls in comprehending HF words in lexical decision tasks administered in 1983 and 1997–1999, but highly significant and exaggerated age-linked declines in comprehending LF words (see also MacKay and James, 2002).

I conclude with claim three of Kensinger et al. (2001) that H.M.'s language processing is unimpaired at grammatical levels. Like the lexical-level data, the syntax-level data of Kensinger et al. cannot discriminate between systems theory vs. binding theory for the same reasons noted earlier in this letter: flawed methodology, the possibility of perfect performance based solely on processing of HF information, and massive repetition of critical stimuli and responses. As an example of flawed methodology, both H.M. and controls in the second syntax comprehension experiment of Kensinger et al. (SC II) scored over 98% correct, a ceiling effect that made deficits impossible to observe. As an example of massive repetition, H.M. and controls in the SC I of Kensinger et al. produced yes–no grammaticality judgments for 128 sentences resembling "yesterday I tied my shoe" (grammatical) and "yesterday I try it on" (ungrammatical), and all 128 trials involved basically similar verb tense cues to the correct response. Massive repetition also characterized the tasks of Kensinger et al. involving production of morphology. For example, a regular past-tense suffix (e.g., *talk–talked*) was the correct response on 64 trials in the past-tense experiment of Kensinger et al. Because binding theory does not predict deficits for familiar and massively repeated information, the results of Kensinger et al. are consistent with binding theory and the results of MacKay et al., in which massive repetition was not a factor. For example, H.M. generated significantly more morphological errors than did age-matched controls for words containing unpredictable suffixes presented without repetition across trials in MacKay and James, 2001, 2002). Similarly, six experiments by MacKay et al. (in press) showed with appropriate procedures that H.M. exhibits *selective* syntax-level comprehension deficits that are consistent with the results of Kensinger et al. *and* with binding theory.

REFERENCES

James LE, MacKay DG. 2001. H.M., word knowledge and aging: Support for a new theory of long-term retrograde amnesia. Psychol Sci 12:485–492.

Kensinger EA, Ullman MT, Corkin S. 2001. Bilateral medial temporal lobe damage does not affect lexical or grammatical processing: Evidence from amnesic patient H.M. Hippocampus 11: 347–360.

Lackner JR. 1974. Observations on the speech processing capabilities of an amnesic patient: Several aspects of H.M.'s language function. Neuropsychologia 12:199–207.

MacKay DG, Burke DM. 1990. Cognition and aging: A theory of new learning and the use of old connections. In: Hess T, editor. Aging and cognition: Knowledge organization and utilization. Amsterdam: North-Holland. p 281–300.

MacKay DG, James LE. 2001. The binding problem for syntax, semantics, and prosody: H.M.'s selective sentence-reading deficits under the theoretical-syndrome approach. Lang Cogn Process 16:419–460.

MacKay DG, James LE. 2002. Aging, retrograde amnesia, and the binding problem for phonology and orthography: A longitudinal study of "hippocampal amnesic" H.M. Aging Neuropsychol Cogn 9:298–333.

MacKay DG, Burke DM, Stewart R. 1998a. H.M.'s language production deficits: Implications for relations between memory, semantic binding, and the hippocampal system. J Mem Lang 38:28–69.

MacKay DG, Stewart R, Burke DM. 1998b. H.M. revisited: Relations between language comprehension, memory, and the hippocampal system. J Cogn Neurosci 10:377–394.

MacKay DG, James LE, Taylor JK, Marian DE. Amnesic H.M. exhibits parallel deficits and sparing in language and memory: Systems versus binding theory accounts. Lang Cogn Process (in press).

Manns JR. 2004. J.F.K., L.B.J., and H.M.: The famous memories of a famous amnesic. Hippocampus 14:411–412.

Scoville WB. 1968. Amnesia after bilateral mesial temporal-lobe excision: Introduction to case H.M. Neuropsychologia 6:211–213.