# Developmental amnesia and its relationship to degree of hippocampal atrophy

E. B. Isaacs*, F. Vargha-Khadem*, K. E. Watkins†, A. Lucas*, M. Mishkin‡§, and D. G. Gadian*

*Institute of Child Health, University College London, London WC1N 1EH, United Kingdom; †Montreal Neurological Institute, McGill University, Montreal, QC, Canada H3A 1S2; and ‡National Institute of Mental Health, Bethesda, MD 20892

Two groups of adolescents, one born preterm and one with a diagnosis of developmental amnesia, were compared with age-matched normal controls on measures of hippocampal volume and memory function. Relative to control values, the preterm group values showed a mean bilateral reduction in hippocampal volume of 8–9% (ranging to 23%), whereas the developmental amnesic group values showed a reduction of 40% (ranging from 27% to 56%). Despite equivalent IQ and immediate memory scores in the two study groups, there were marked differences between them on a wide variety of verbal and visual delayed memory tasks. Consistent with their diagnosis, the developmental amnesic group was impaired relative to both other groups on nearly all delayed memory measures. The preterm group, by contrast, was significantly impaired relative to the controls on only a few memory measures, i.e., route following and prospective memory. We suggest that early hippocampal pathology leads to the disabling memory impairments associated with developmental amnesia when the volume of this structure is reduced below normal by ≈20–30% on each side. Whether this is a sufficient condition for the disorder or whether abnormality in other brain regions is also necessary remains to be determined.

**D**evelopmental amnesia is a selective disorder characterized by marked impairment in episodic memory despite relatively preserved semantic memory (1, 2). The disorder is associated with bilateral medial temporal pathology that seems to be restricted mainly to the hippocampus, with some involvement of the putamen, thalamus, and right retrosplenial cortex demonstrated in voxel-based morphometry studies (2–4). Because of their frequent failure to remember the events of everyday life, children with developmental amnesia are seriously disabled, yet they have managed to acquire literacy skills and factual knowledge in line with a level of intellectual ability that ranges from low to high average. The disorder has been observed in children with pathology incurred at ages extending from birth to puberty (3, 4), with a frequent etiology of one or more hypoxic/ischemic episodes.

An important question raised by these findings concerns what degree of damage gives rise to developmental amnesia. Because medial temporal pathology seems to play a critical role in the disorder, and because the clearest indicator in the cases studied is hippocampal atrophy, the question becomes one of determining how much hippocampal atrophy is necessary for the condition to emerge.

Here we compare a group of children with developmental amnesia (DA group) with another group that has a significant reduction in hippocampal volume and yet does not exhibit the same disabling memory impairments (5). The latter group consisted of children born preterm with very low birth weight (PT group). Direct comparison between the two groups, in terms of both their memory abilities and hippocampal volumes, could help to determine the amount of damage that causes developmental amnesia and, more generally, the mnemonic deficits produced by different degrees of hippocampal atrophy.

## Methods

**Subjects.** The DA group consisted of six children who had suffered hypoxic/ischemic episodes either at birth or within the first year (yr) of life and four others who had sustained such episodes between the ages of 6 and 14 yr. These two subgroups were found to have largely indistinguishable profiles both behaviorally and neuropathologically (3, 4). For the purpose of the present investigation they were therefore combined into a single group of 10 (six male, four female).

The PT group consisted of 11 (eight male, three female) children delivered at <30 weeks gestation (median of 28 weeks). These children, as described by Isaacs *et al.* (5), all had very low birth weight (<1,500 g) but were neurologically normal. Although they had required at least 7 days of ventilation (median of 9 days), this was not regarded as evidence of a history of hypoxia, and they were not included on this basis. They were selected for comparison with the control group (C group) because they were known to have reduced hippocampal volumes and yet did not exhibit the same disabling memory impairments as the DA group. Their inclusion allowed a wider range of hippocampal volume reductions to be examined in relation to memory impairment.

The C group consisted of eight (three male, five female) normally developing adolescents, born at full term. Mean ages at assessment for the groups were 14 yr 1 mo (DA), 13 yr 7 mo (PT), and 13 yr 8 mo (C).

The study was approved by the Great Ormond Street Hospital for Children/Institute of Child Health Research Ethics Committee, as well as by the appropriate regional ethics committees, and informed consent was obtained from each of the patients and control subjects.

**Neuroimaging.** Structural MRI studies were performed on a 1.5-T Siemens Vision system (Erlangen, Germany). T1- and T2-weighted scans were acquired: (*i*) magnetization-prepared rapid acquisition gradient echo (MPRAGE) 3D T1-weighted volume acquisition (6) with repetition time of 10 ms, echo time of 4 ms, inversion time of 200 ms, flip angle of 12°, matrix size of 256 × 256, field of view equal to 250 mm, partition thickness of 1.25 mm, 128 sagittal partitions in the third dimension, and acquisition time of 8.3 min; and (*ii*) coronal and axial turbo spin-echo T2-weighted scans with repetition time of 4,600 ms, echo time of 90 ms, and acquisition time of 4.3 min for each orientation.

Hippocampal abnormalities were assessed quantitatively by measuring hippocampal volumes (by one observer, D.G.G.). The 3D data sets were reformatted to 1-mm-thick contiguous slices in a tilted coronal plane perpendicular to the long axis of the hippocampus. Cross-sectional areas were measured along the entire length of the hippocampus by using every third slice, as described previously (7). The volumes were calculated by summing the cross-sectional areas and multiplying them by the distance between the measured slices (i.e., 3 mm). Intracranial volumes were measured from the unreformatted sagittal 3D data

---

**Table 1. Hippocampal volumes for the three groups**

| Group | n | LHCV | | RHCV | | % of C group, right | | % of C group, left | |
|---|---|---|---|---|---|---|---|---|---|
| | | Mean | SD | Mean | SD | Mean | Range | Mean | Range |
| DA | 10 | 2,269 | 362 | 2,367 | 326 | 59 | 44–73 | 60 | 45–70 |
| PT | 11 | 3,546 | 295 | 3,652 | 366 | 91 | 77–103 | 92 | 75–105 |
| C | 8 | 3,883 | 251 | 3,980 | 260 | NA | NA | NA | NA |

Mean left and right volumes (in mm$^3$) and volumes expressed as a percentage of the C group mean are given. LHCV, left hippocampal volume; RHCV, right hippocampal volume.

sets (7). The hippocampal volumes then were corrected for intracranial volume (7) and are presented here in this corrected form.

**Neuropsychology.** *Intelligence and academic attainments.* To establish the level of general intellectual ability achieved by the three groups, the Wechsler Intelligence Scale for Children, Third Edition (WISC-III) (8) was administered in full, allowing calculation of verbal, performance, and full-scale IQs. The information, vocabulary, and comprehension subtests of the WISC-III also provided information about semantic memory ability.

Literacy and numeracy skills were measured by the Wechsler Objective Reading Dimensions (WORD) (9) and Wechsler Objective Numerical Dimensions (WOND) (10) tests. The three subtests of the WORD (basic reading, spelling, and reading comprehension) and the two subtests of the WOND (numerical operations and mathematical reasoning) tests each provided standard scores with a mean of 100 and a SD of 15.

For each child, the WORD and WOND scores predicted by their WISC-III IQ scores were compared with the scores they actually obtained to determine whether the two were significantly different.

*Memory.* A series of commonly used laboratory tests considered to be measures of episodic memory was used to measure verbal and visual recall in both immediate and delayed conditions. The Wechsler Memory Scale, Form I (11) was administered in the standard way, but with age corrections for children under the age of 18 yr. Measures of immediate and delayed recall are obtained for (*i*) story recall, in which two prose passages are read for immediate and unalerted recall after a 90-min delay, and (*ii*) paired-associate learning, in which 10 pairs of words, six semantically related (e.g., up and down) and four unrelated (e.g., cabbage and pen), are read three times for immediate recall after each presentation and delayed recall after a 90-min delay. The Children's Auditory Verbal Learning Test 2 (CAVLT-2) (12), provides measures of immediate memory, level of learning, and interference effects, as well as delayed recall and recognition of verbal material. In this test, a list of 16 words is read to the child five times, with free recall after each presentation, and, to determine the effects of interference, a second 16-word list is read once for immediate recall. After an interval of 20 min, delayed recall of the original list is obtained, followed by forced choice (i.e., yes or no) recognition of the items on this list. The design learning subtest of the Adult Memory and Information Processing Battery (13) was administered as a visual, nonverbal analogue of the CAVLT-2. In this design test, which yields measures similar to those of the CAVLT-2, a pattern, formed by joining dots in a matrix, is presented for five learning trials, followed by an interference trial and immediate and delayed recall. Finally, the Rey–Osterrieth Complex Figure (14) was presented, first for copying and then for reproduction from memory after a 40-min interval.

In addition to the above tasks, the Rivermead Behavioural Memory Test (15), designed to measure recognition and recall of events closer to real-life situations, was used to quantify everyday memory. A Standardized Profile Score was obtained by assigning a score of 0, 1, or 2 to each of 11 subtests: immediate and delayed memory for a name, face, picture, story, and route, and measures of prospective memory (appointment, message, and belonging), plus orientation to time and place. A 12th subtest, knowledge of the date, was not given to some of the younger amnesic cases and was therefore omitted from the analysis, reducing the maximum Standardized Profile Score from 24 to 22. Three composite scores were created by summing the scores for (*i*) story, immediate and delayed recall; (*ii*) route, immediate and delayed recall; and (*iii*) prospective memory, appointment, message, and belonging.

**Statistics.** Because the groups were relatively small and the distributions of some variables were skewed, group differences were evaluated with nonparametric statistics. Differences among all three groups were assessed with the Kruskal–Wallis test, whereas paired comparisons were made either with the Mann–Whitney $U$ test (unrelated samples) or the Wilcoxon Signed Ranks test (related samples). Correlations were calculated by using Spearman's coefficient.

## Results

**Hippocampal Volumes.** Group means for left and right hippocampal volumes, measured in cubic millimeters, are presented in Table 1. There was a highly significant difference among groups for both left ($\chi^2 = 21.2, P < 0.0001$) and right ($\chi^2 = 19.7, P < 0.0001$) hippocampi. Pairwise comparisons showed a graded pattern: both hippocampi were smaller in the PT than in the C group (left: $U = 17, P < 0.03$; right: $U = 19, P < 0.04$) and smaller in the DA than in the PT group (left: $U = 0, P < 0.0001$; right: $U = 1, P < 0.0001$). A comparison of the two hippocampi within each group indicated that the right hippocampus was larger than the left in all three groups, but the difference reached significance only in the C group ($Z = -2.0, P < 0.05$).

Table 1 also shows the mean left and right hippocampal volumes (LHCV and RHCV, respectively) for the two study groups calculated as percentages of the mean volumes of the C group. The average bilateral reduction below C group values in the PT group was 8.5%, whereas in the DA group it was 40.5%, with no overlap between the two (range of $-3\%$ to 23% and 27–56%, respectively).

**Intelligence and Academic Attainments.** There were significant differences among the three groups for all three IQ indices: verbal IQ (VIQ), $\chi^2 = 6.9, P < 0.03$; performance IQ (PIQ), $\chi^2 = 6.2, P < 0.04$; and full-scale IQ (FSIQ), $\chi^2 = 8.7, P < 0.01$ (Fig. 1). These differences were attributable to significantly lower scores in both the PT and DA groups than in the C group for VIQ (PT: $U = 16.5, P < 0.02$; DA: $U = 14, P < 0.02$) and FSIQ (PT: $U = 14, P < 0.02$; DA: $U = 12, P < 0.01$), and in the DA group compared with the C group for PIQ ($U = 15; P < 0.03$). The two experimental groups did not differ significantly from each other, nor did VIQ and PIQ differ within any of the groups. Although the PT and DA groups were impaired, their mean scores still fell within the average range described by Wechsler
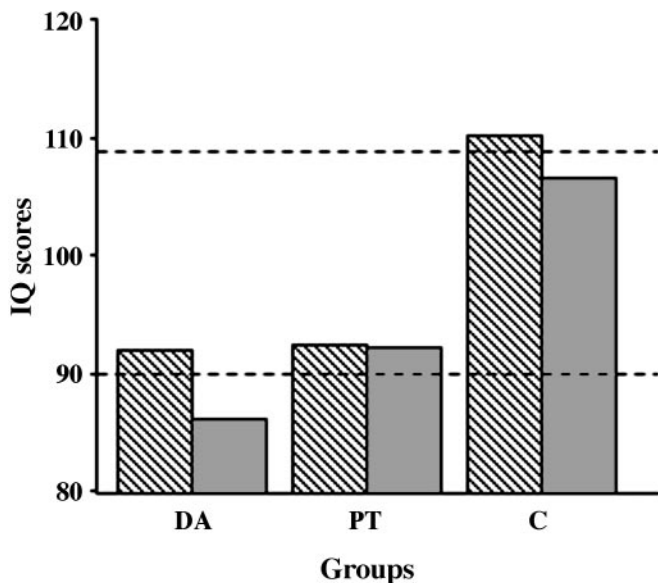
NEUROSCIENCE

**Fig. 1.** Mean scores for verbal IQ (hatched bars) and performance IQ (filled bars) for the DA, PT, and C groups. Dotted lines denote the average range.

(8), i.e., 90–109; they differed from the C group in part because the latter obtained a mean FSIQ of 110.

Analyses of the Wechsler Intelligence Scale for Children, Third Edition subtest scores revealed differences among the three groups only for arithmetic ($\chi^2 = 12.4, P < 0.002$), digit span ($\chi^2 = 9.5, P < 0.009$), and symbol search ($\chi^2 = 6.9, P < 0.03$). Paired comparisons performed to determine the source of these differences indicated that the PT and DA groups scored significantly below the C group on arithmetic (PT: $U = 10.5, P < 0.004$; DA: $U = 2.5, P < 0.001$), whereas only the PT group differed from the C group on digit span ($U = 9, P < 0.003$), and only the DA group did so on symbol search ($U = 11, P < 0.03$). The DA and PT groups did not differ from each other on any of the subtests including, in particular, information, vocabulary, and comprehension, all measures of semantic knowledge (see Table 2, which is published as supporting information on the PNAS web site, www.pnas.org).

Analyses of the Wechsler Objective Reading Dimensions and Wechsler Objective Numerical Dimensions subtest scores revealed significant differences among the groups by all measures, except basic reading (spelling: $\chi^2 = 7.6, P < 0.02$; reading comprehension: $\chi^2 = 6.8, P < 0.03$; numerical operations: $\chi^2 = 14.5, P < 0.001$; mathematical reasoning: $\chi^2 = 7.2, P < 0.03$). Pairwise comparisons indicated that the DA group scored below the C group on spelling ($U = 7.5, P < 0.002$), reading comprehension ($U = 12.5, P < 0.01$), numerical operations ($U = 9, P < 0.03$), and mathematical reasoning ($U = 10, P < 0.02$), whereas the PT group scored below the C group on numerical operations ($U = 1.5, P < 0.0001$) and mathematical reasoning ($U = 16.5, P < 0.02$). Despite these differences, the mean scores for the two experimental groups, which did not differ significantly from each other, all fell within one SD of the normal mean, in line with their IQs. Numerical operations was the only attainment test on which the PT group obtained the lowest mean group score, significantly below the DA group ($U = 17.0, P < 0.05$), thus confirming the observation that these children have a high incidence of calculation difficulties (16).

In summary, although they were poorer than the C group by many of the above measures, the DA and PT groups did not differ significantly from each other by any of these measures and

demonstrated levels of intelligence, literacy, and numeracy that fall within the low-average to average range.

**Memory.** Kruskal–Wallis tests carried out on the scores of the three groups on the verbal and visual memory tests revealed highly significant group effects by most of the measures (see Table 3, which is published as supporting information on the PNAS web site). The only exceptions were two measures of immediate visual learning. Examination showed that, with the exception of the two digit-span measures (see below), the mean scores always were graded in the order C > PT > DA, although not always significantly different. Where appropriate, paired comparisons indicated that, by most of these measures, the DA group was significantly impaired relative to each of the other groups, which did not differ from each other (C = PT > DA). In contrast to the DA group, the PT group was impaired relative to the C group by two measures only: overall memory quotient and immediate story recall. The PT group also did poorly on the digit-span measures, particularly on backward recall, where their mean score was significantly lower than that of both other groups (C: $U = 9.0, P < 0.003$; DA: $U = 27.0, P < 0.04$).

The Rivermead Behavioural Memory Test yielded a pattern of results similar to that found with most of the other memory tests; i.e., overall group differences were significant for the Standardized Profile Score and scores on all subtests except orientation and face recognition, and, when significant group differences occurred, the DA group always performed significantly more poorly than the C group (see Table 4, which is published as supporting information on the PNAS web site). In most instances, the PT group did not differ from the C group, but there were exceptions. Thus, on the Standardized Profile Score, delayed route recall, and prospective memory composite score, the PT group scored significantly below the C group but significantly above the DA group. Finally, the PT group's performance was depressed to a level that was not significantly different from that of the DA group on the route composite and the message. No child in the PT group presented with an amnesic profile.

Correlation analyses were carried out within each group to examine whether there was any within-group relationship between hippocampal volume and any of the memory scores. No significant relationships were found.

**Discussion**

The main aim of our study was to determine the degree of bilateral hippocampal atrophy that gives rise to developmental amnesia. In the 10 children with this disorder, all of whom presented with memory impairments after earlier hypoxic/ischemic episodes, the bilateral volume reduction of the hippocampus relative to control values ranged from 27% to 56%. Our findings suggest that this degree of atrophy is a prerequisite for the memory impairments shown by these children; if lesser damage led to the memory impairment, it is likely that the DA group would have included individuals with less damage. By way of comparison, none of the children in the PT group showed the pattern of memory deficits exhibited by each of the children in the DA group, nor was there any overlap in hippocampal volumes between the two groups. Although differing extrahippocampal abnormalities arising from the different etiologies may have had a crucial influence on the differing memory profiles of the two groups (see below), our findings nevertheless enable us to suggest that the hippocampal volume reduction required for the symptoms of developmental amnesia to emerge is $\approx 20$–$30\%$ on each side. [Of interest in this connection is evidence that the amount of hippocampal damage needed to produce a delayed recognition impairment in monkeys is also $\approx 20\%$ (17)]. That this amount of damage must be bilateral is indicated by the fact that unilateral removal of most of the hippocampus during left or

right temporal lobe surgery for relief of epilepsy does not produce amnesia (18). Further information about the precise relationship between degree of atrophy and extent of memory impairment will clearly emerge if we find individuals with lesser degrees of bilateral atrophy associated with hypoxic/ischemic episodes. Such individuals would help to establish to what extent developmental amnesia associated with hippocampal necrosis is a graded or threshold phenomenon.

The proposal that bilateral hippocampal atrophy on the order of 20–30% may be necessary for the disorder to appear does not imply that such damage is sufficient. Additional pathology in other brain structures also may be necessary, either that already detected in the DA group, such as in the putamen and posterior thalamus bilaterally and in the retrosplenial area on the right (2–4), or that still undetected. Furthermore, we cannot preclude the possibility that the degree of hippocampal atrophy underestimates the ''functional'' lesion, although studies of neurotoxic lesions in monkeys suggest that, for cell loss of 40% or less, the hippocampal volume reduction measured by MRI underestimates the total cell loss, as determined histologically, by only ≈5% on average (19).

Although the PT group showed few deficits on measures of episodic memory, they did fall into the poor memory category on the Rivermead Behavioural Memory Test, notably because nearly all of the children in this group failed in retracing a route around the room and in recalling cues they were to respond to later (the prospective memory composite). A mean reduction of 8–9% in hippocampal volume apparently was sufficient to disrupt memory of these items, suggesting that these particular items may be tapping basic functions of the hippocampus.

The DA group was far more severely impaired than the PT group by all of the measures of episodic memory, and yet there were no differences between the two in level of either intelligence or academic attainments. Ontogenetic development of these cognitive abilities thus seems to be largely independent of the hippocampus. As suggested earlier (1, 20), the substrate for long-term storage of the semantic information required for their development thus appears to be dissociable from that for episodic memory and may depend instead on the cortex of the parahippocampal region.

Determining the extent of hippocampal volume reduction that leads to developmental amnesia is not just of theoretical interest. The disorder tends to remain undiagnosed until children are of school age, when the increased requirement for carrying out activities independently begins to highlight their memory difficulties. Conducting hippocampal volumetric studies in children who have recently suffered a hypoxic/ischemic episode could help to identify those in an at-risk group because of bilateral reductions exceeding 20–30%.

1. Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W. & Mishkin, M. (1997) *Science* **277,** 376–380.
2. Gadian, D. G., Aicardi, J., Watkins, K. E., Porter, D. A., Mishkin, M. & Vargha-Khadem, F. (2000) *Brain* **123,** 499–507.
3. Vargha-Khadem, F., Gadian, D. G. & Mishkin, M. (2001) *Philos. Trans. R. Soc. London B* **356,** 1435–1440.
4. Vargha-Khadem, F., Salmond, C. H., Watkins, K. E., Friston, K. J., Gadian, D. G. & Mishkin, M. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 10055–10060.
5. Isaacs, E. B., Lucas, A., Chong, W. K., Wood, S. J., Johnson, C. L., Marshall, C., Vargha-Khadem, F. & Gadian, D. G. (2000) *Pediatr. Res.* **47,** 713–720.
6. Mugler, J. P. & Brookeman, J. R. (1990) *Magn. Reson. Med.* **15,** 152–157.
7. Van Paesschen, W., Connelly, A., King, M. D., Jackson, G. D. & Duncan, J. S. (1997) *Ann. Neurol.* **41,** 41–51.
8. Bracken, B. A., ed. (1992) *Wechsler Intelligence Scale for Children* (Psychological Corp., Sidcup, Kent, U.K.), 3rd U.K. Ed.
9. Rust, J., Golombok, S. & Trickey, G. (1993) *Wechsler Objective Reading Dimensions* (Psychological Corp., Sidcup, Kent, U.K.).
10. Rust, J. (1996) *Wechsler Objective Numerical Dimensions* (Psychological Corp., Sidcup, Kent, U.K.).
11. Wechsler, D. & Stone, C. P. (1945) *Wechsler Memory Scale* (Psychological Corp., San Antonio, TX).
12. Talley, J. L. (1993) *Children's Auditory Verbal Learning Test-2 (CAVLT-2)* (Psychological Assessment Resources, Odessa, FL).
13. Coughlan, J. & Hollows, S. E. (1985) *Adult Memory and Information Processing Battery* (St. James Hospital, Leeds, U.K.).
14. Rey, A. (1964) *L'Examen Clinique en Psychologie* (Presses Universitaire de France, Paris).
15. Wilson, B., Cockburn, J. & Baddeley, A. (1991) *The Rivermead Behavioural Memory Test* (Thames Valley Test Co., Bury St. Edmunds, Suffolk, U.K.), 2nd Ed.
16. Isaacs, E. B., Edmonds, C. J., Lucas, A. & Gadian, D. G. (2001) *Brain* **124,** 1701–1707.
17. Zola, S. M. & Squire, L. R. (2001) *Hippocampus* **11,** 92–98.
18. Milner, B. (1970) in *Biology of Memory*, eds. Pribram, K. H. & Broadbent, D. E. (Academic, New York).
19. Malkova, L., Lex, C. K., Mishkin, M. & Saunders, R. C. (2001) *Hippocampus* **11,** 361–370.
20. Mishkin, M., Suzuki, W., Gadian, D. G. & Vargha-Khadem, F. (1997) *Philos. Trans. R. Soc. London B* **352,** 1461–1467.

NEUROSCIENCE