LETTERS
https://doi.org/10.1038/s43587-021-00117-4

nature aging

Check for updates

Dissociable effects of *APOE* ε4 and β-amyloid pathology on visual working memory

Kirsty Lu[1]✉, Jennifer M. Nicholas[2], Yoni Pertzov[3], John Grogan[4], Masud Husain[4,5], Ivanna M. Pavisic[1,6], Sarah-Naomi James[7], Thomas D. Parker[1], Christopher A. Lane[1], Ashvini Keshavan[1], Sarah E. Keuss[1], Sarah M. Buchanan[1], Heidi Murray-Smith[1], David M. Cash[1,6], Ian B. Malone[1], Carole H. Sudre[1,7,8,9], William Coath[1], Andrew Wong[7], Susie M. D. Henley[1], Nick C. Fox[1,6], Marcus Richards[7], Jonathan M. Schott[1,10] and Sebastian J. Crutch[1,10]

Although *APOE* ε4 carriers are at substantially higher risk of developing Alzheimer's disease than noncarriers[1], controversial evidence suggests that *APOE* ε4 might confer some advantages, explaining the survival of this gene (antagonistic pleiotropy)[2,3]. In a population-based cohort born in one week in 1946 (assessed aged 69–71 years), we assessed differential effects of *APOE* ε4 and β-amyloid pathology (quantified using ^{18}F-Florbetapir-PET) on visual working memory (object–location binding). In 398 cognitively normal participants, *APOE* ε4 and β-amyloid had opposing effects on object identification, predicting better and poorer recall, respectively. ε4 carriers also recalled locations more precisely, with a greater advantage at higher β-amyloid burden. These results provide evidence of superior visual working memory in ε4 carriers, showing that some benefits of this genotype are demonstrable in older age, even in the preclinical stages of Alzheimer's disease.

The human apolipoprotein E (*APOE*) gene has three main alleles—ε2, ε3 and ε4—with ε3 being the most common. *APOE* ε4 is the strongest genetic risk factor for sporadic Alzheimer's disease, increasing risk in a dose-dependent manner (three- to fourfold for ε4 heterozygotes;12–16-fold for ε4 homozygotes[1,2]) and resulting in earlier dementia onset[1]. The role of apoE in Alzheimer's disease pathogenesis is not fully understood, but ε4 is associated with reduced clearance and hence accumulation of β-amyloid[1,4]. β-amyloid pathology is detectable up to two to three decades before symptoms emerge[5,6], is associated with subtle cognitive deficits in cognitively unimpaired older adults[7] and is necessary, but not sufficient, for a diagnosis of Alzheimer's dementia[8].

However, a diverse range of survival advantages have been reported in ε4 carriers including increased fertility[2,9,10], resistance to infections[2], decreased perinatal and infant mortality[2] and some slight cognitive advantages[2,3,11–17]. *APOE* ε4 may be an example of antagonistic pleiotropy—a leading evolutionary explanation for aging and noncommunicable disease[9,18]—whereby a gene has both beneficial and detrimental effects, with the latter generally manifesting later in life when the forces of natural selection are weaker[18]. This hypothesis may explain why *APOE* ε4 (the ancestral allele) persists in human populations rather than being replaced by the ε3 and ε2 alleles, which evolved later[2,19,20]; however, evidence for its putative cognitive benefits remains mixed and controversial[3,13,21].

One cognitive measure where ε4 carriers have shown superior performance is the 'What was where?' visual working memory test[22]: ε4 carriers have been reported to recall object locations more accurately than noncarriers after delays of a few seconds[23–25]. These studies did not, however, evaluate the possible influence of preclinical Alzheimer's disease pathology. One notable feature of this task is its analog measure of location memory (in contrast to traditional 'correct or incorrect' measures), which allows fine-grained assessment of the precision or quality of memory representations[26,27].

We aimed to assess the relative influence of *APOE* ε4 and β-amyloid pathology on the 'What was where?' task in a large population-based sample of adults from Insight 46, a substudy of the MRC National Survey of Health and Development (the UK 1946 Birth Cohort)[28]—the world's longest continuously running birth cohort[28,29]. Participants were aged ~70 years—an age when rates of dementia are low but a substantial proportion (~15–25%) have biomarker evidence of preclinical Alzheimer's disease[30,31]. Based on the literature, we hypothesized that *APOE* ε4 would be associated with slightly more accurate recall of object locations but that β-amyloid pathology would be associated with subtly poorer performance across the task. Because these two predicted effects are in opposition for ε4 carriers with elevated amyloid burden, we aimed to explore interactions between *APOE* ε4 and β-amyloid on visual working memory. We also aimed to test whether this task is sensitive to differences in hippocampal volume and white matter hyperintensity volume (WMHV, a marker of cerebral small vessel disease that is common in older people and is associated with cognitive decline, particularly in executive function[32]).

Results

Participants underwent *APOE* genotyping, amyloid positron emission tomography/magnetic resonance imaging (PET/MRI)

[1]Dementia Research Centre, Department of Neurodegenerative Disease, UCL Queen Square Institute of Neurology, University College London, London, UK. [2]Department of Medical Statistics, London School of Hygiene and Tropical Medicine, London, UK. [3]Department of Psychology, The Hebrew University of Jerusalem, Jerusalem, Israel. [4]Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK. [5]Department of Experimental Psychology, University of Oxford, Oxford, UK. [6]UK Dementia Research Institute at UCL, University College London, London, UK. [7]MRC Unit for Lifelong Health and Ageing at UCL, University College London, London, UK. [8]Centre for Medical Image Computing, Department of Computer Science, University College London, London, UK. [9]School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK. [10]These authors jointly supervised this work: Jonathan M. Schott, Sebastian J. Crutch. ✉e-mail: kirsty.lu@ucl.ac.uk

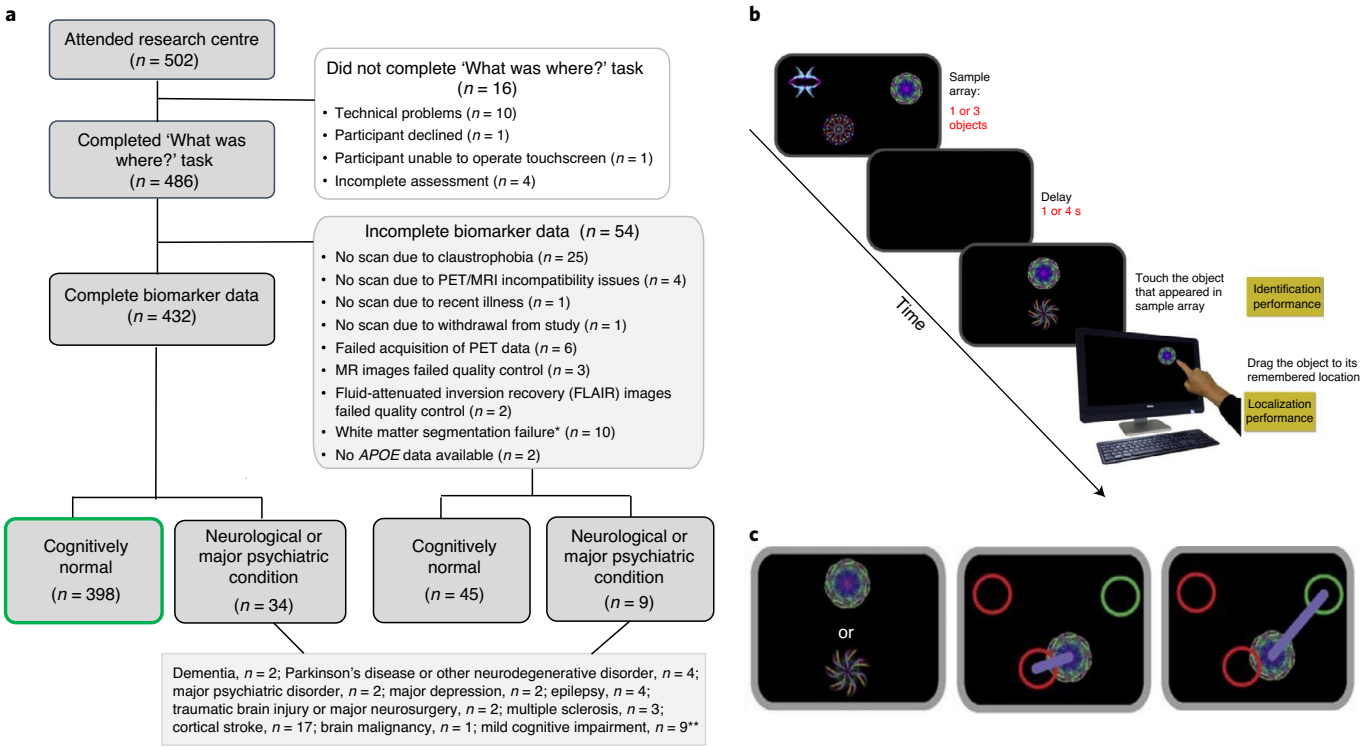

Fig. 1 | Study design. a, Flow-chart of data acquisition and reasons for missing data. A total of 398 cognitively normal participants with complete biomarker data formed the main analysis sample. See Methods and ref. [42] for further details on the definitions of neurological and major psychiatric disorders. *In most cases this was due to erroneous segmentation of vascular abnormalities such as stroke or demyelination. **The numbers in this box add up to 46 because some participants had more than one condition. **b**, Presentation of the 'What was where?' task. **c**, Illustration of outcome measures. Green circles indicate the original location of the target object, red circles indicate the original locations of nontarget objects and blue lines indicate measured localization error. Left, object identification; the participant is required to select the object that they remember seeing. Middle, localization error is measured from the location reported by the participant to that of the closest object; if the reported location is within 140 pixels of that of a nontarget object, this is considered a misbinding error. Right, gross localization error is measured from the location reported by the participant to the original location of the target object. Credit: **b**,**c**, are reprinted from Liang et al.[42] under the terms of the Creative Commons Attribution License (CC BY): https://creativecommons.org/licenses/by/4.0/.

neuroimaging and neuropsychological assessment including the 'What was where?' visual working memory task (Methods; Fig. 1b,c). A total of 486 participants completed the task: 398 were cognitively normal, with complete biomarker data (Fig. 1a), of whom 120 (30%) were *APOE* ε4 carriers. Participant characteristics are provided in Supplementary Table 1, along with descriptive statistics for the primary outcome measures. Performance on established tests of memory is presented in Supplementary Information. The prevalence of amyloid positivity among ε4 carriers and non-ε4 carriers was 37.5 and 9.7%, respectively ($\chi^2 = 43.7$, $P < 0.0001$), consistent with the existing literature[30].

In cognitively normal participants with complete biomarker data, multivariable regression models were fitted (Methods) to investigate associations between task performance and *APOE* ε4 (carrier or noncarrier), amyloid status (positive or negative), hippocampal volume and WMHV. The models also included the task condition factors of memory load (low or high) and delay interval (short or long), as well as adjusting for head size and demographic and life-course factors, previously shown to predict cognitive performance throughout adulthood in this cohort (Methods). Where between-individual factors were significantly associated with performance, we tested for interactions with delay to investigate whether group differences were due to better retention over time. We also tested for interactions between *APOE* ε4 and amyloid status, to investigate whether the effects of *APOE* ε4 differed between amyloid-positive and -negative groups.

Results of the regression models are given in Table 1 (see Supplementary Information for results relating to demographic and life-course factors). As expected, identification and localization memory were poorer under the high-load than low-load condition, and localization was also poorer after long compared to short delay (Table 1). However, in contrast to previous studies[22,23,33], delay had no statistically significant effect on the proportion of identification or misbinding errors (Table 1).

Identification. On average, amyloid-positive participants were 19% more likely to make identification errors than amyloid-negative participants ($P = 0.029$; adjusted error rate (95% CI): amyloid-positive 0.20 (0.17–0.22); amyloid-negative 0.17 (0.16–0.18)) (Table 1 and Fig. 2a). Independently, *APOE* ε4 carriers were 14% less likely to make identification errors ($P = 0.026$; ε4 carriers 0.16 (0.15–0.17); non-ε4 carriers 0.18 (0.17–0.19)) (Table 1 and Fig. 2a). These group differences in error rates are very small in magnitude: error rates of 0.16 and 0.18 equate to 3.8 and 4.3 errors, respectively. There was no evidence of a statistically significant interaction between amyloid status and *APOE* ε4 for identification errors (OR = 0.88 (0.63–1.15), $P = 0.29$). Similar dissociable effects of amyloid status and *APOE* ε4 were observed on a verbal story recall task (Supplementary Information).

The association between *APOE* ε4 and identification was consistent across long and short delays (OR for interaction between *APOE* ε4 and delay = 0.97 (0.77–1.22), $P = 0.79$), as was that between

Table 1 | Associations between biomarkers and 'What was where?' outcomes in cognitively normal participants ($n = 398$)

Predictor	Primary outcome measures			2D-mixture model outcomes		
	Identification errors: OR for incorrect response (95% CI)	Localization error coefficient (95% CI)[a]	Misbinding errors: OR for error (95% CI)	Imprecision parameter (pixels): coefficient (95% CI)	Misbinding parameter: OR (95% CI)[b]	Guessing parameter: OR (95% CI)[c]
β-Amyloid status (negative as reference)	1.19* (1.02, 1.38) $P = 0.029$	1.02 (0.95, 1.10) $P = 0.58$	0.94 (0.77, 1.15) $P = 0.55$	3 (−4, 9) $P = 0.42$	0.96 (0.78, 1.20) $P = 0.74$	1.03 (0.57, 1.88) $P = 0.92$
White matter hyperintensity volume (per 10 ml)	0.98 (0.89, 1.09) $P = 0.77$	0.99 (0.95, 1.03) $P = 0.67$	0.97 (0.88, 1.08) $P = 0.63$	0 (−4, 5) $P = 0.84$	0.98 (0.86, 1.11) $P = 0.73$	0.89 (0.64, 1.25) $P = 0.50$
Hippocampal volume (per ml)	0.98 (0.88, 1.09) $P = 0.69$	0.98 (0.95, 1.02) $P = 0.41$	1.05 (0.93, 1.19) $P = 0.41$	−3 (−6, 1) $P = 0.19$	1.05 (0.92, 1.20) $P = 0.49$	0.93 (0.66, 1.30) $P = 0.65$
APOE ε4 (noncarriers as reference)	0.86* (0.75, 0.98) $P = 0.026$	0.93** (0.88, 0.98) $P = 0.007$	1.01 (0.87, 1.17) $P = 0.94$	−6* (−11, −1) $P = 0.027$	0.96 (0.82, 1.13) $P = 0.63$	0.75 (0.45, 1.26) $P = 0.28$
High load (low load as reference)	4.42** (3.52, 5.56) $P < 0.001$	1.48** (1.42, 1.54) $P < 0.001$	N/A	N/A	N/A	N/A
Long delay (short delay as reference)	1.05 (0.95, 1.17) $P = 0.32$	1.11** (1.08, 1.14) $P < 0.001$	0.98 (0.87, 1.10) $P = 0.73$	13** (9, 16) $P < 0.001$	0.91 (0.81, 1.03) $P = 0.13$	0.93 (0.66, 1.30) $P = 0.66$

*Significant at $P < 0.05$; **significant at $P < 0.01$. Because these results were obtained from multivariable regression models (Methods), each association is independent of all others. No adjustments were made for multiple comparisons. In addition to the predictors listed, models also adjusted for sex, age at assessment, childhood cognitive ability, education, socioeconomic position and total intracranial volume. For details on the associations of these demographic and life-course factors with visual working memory performance, see Supplementary Information and Supplementary Table 4. [a]Because localization error data were log-transformed, the coefficients are quoted in exponentiated form for ease of interpretation; for example, a coefficient of 1.10 would mean that the predictor was associated with 10% greater localization error while a coefficient of 0.90 would mean that the predictor was associated with 10% smaller localization error. [b]OR > 1 indicates that the predictor is associated with a higher proportion of misbinds. [c]OR > 1 indicates that the predictor is associated with a higher proportion of guesses. CI, confidence interval; N/A, not applicable.

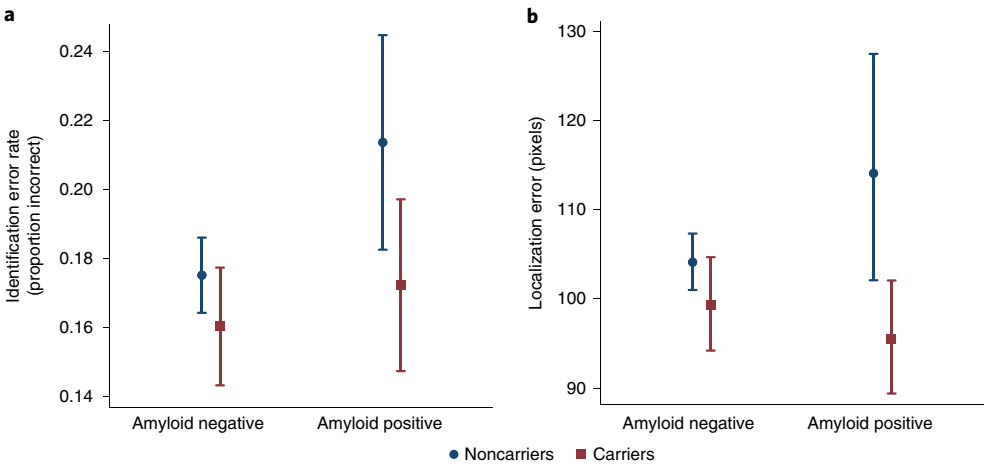

Fig. 2 | Performance on the 'What was where?' task in cognitively normal participants ($n = 398$). a, The main effects of amyloid status and *APOE* ε4 on identification error rate. **b**, The main effects of *APOE* ε4 on localization error. Markers show adjusted means from the multivariable regression models, adjusted for delay (long versus short), load (low versus high), sex, age at assessment, childhood cognitive ability, education, socioeconomic position, white matter hyperintensity volume, hippocampal volume and total intracranial volume. Error bars show 95% CI. Values plotted are essentially unchanged even if the model does not include adjustment for white matter hyperintensity volume, hippocampal volume and total intracranial volume. **b**, Data were log-transformed for analysis, but the means and CI presented here have been back-transformed for ease of interpretation. For numbers of participants in each group, see Supplementary Table 1.

amyloid positivity and identification (OR = 0.83 (0.62–1.09), $P = 0.18$). When assessed as a continuous variable (standardized uptake value ratio (SUVR)), the association between amyloid burden and identification error was not statistically significant although it was in the expected direction (OR = 1.05 (95% CI 0.96–1.15) per 0.1 SUVR increment, $P = 0.26$).

Hippocampal volume and WMHV did not show statistically significant associations with identification (Table 1).

Localization. *APOE* ε4 carriers performed better with respect to spatial memory, on average positioning objects 7% closer to the true location than non-ε4 carriers ($P = 0.007$) (Table 1 and Fig. 2b). In Supplementary Information it is shown that ε4 carriers had a lower mean localization error in 19 out of 24 trials.

There was no evidence of statistically significant difference in localization error between the amyloid groups (Table 1 and Fig. 2b), nor of an association between continuous amyloid burden and localization error (coefficient = 1.00 (0.97–1.04) per 0.1 SUVR increment, $P = 0.90$). Regarding an interaction between amyloid status and *APOE* ε4, the localization memory advantage associated with *APOE* ε4 was greater among amyloid-positive than amyloid-negative participants, but this interaction was not statistically

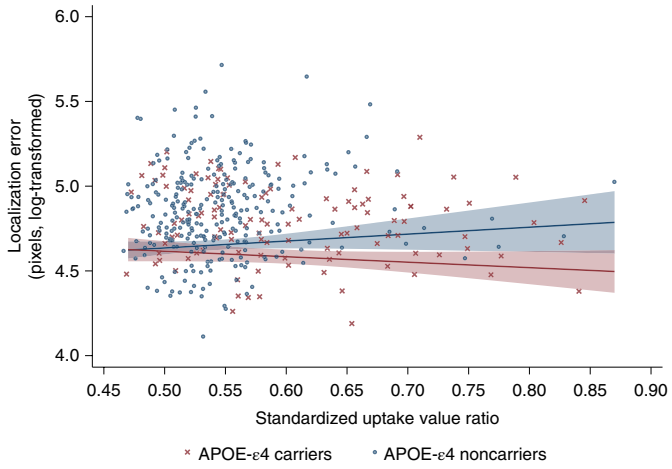

Fig. 3 | Association between β-amyloid burden (quantified using SUVR) and localization error on the 'What was where?' task for *APOE* ε4 carriers ($n=120$) and noncarriers ($n=278$). Solid lines represent marginal means from the multivariable regression model (Methods) and shaded areas represent 95% CI, with ε4 carriers shown in red and noncarriers in blue. Models were adjusted for load (low versus high), delay (short versus long), sex, age, education, socioeconomic position, white matter hyperintensity volume, hippocampal volume and total intracranial volume; no adjustments were made for multiple comparisons. Markers show each participant's mean localization error across the experiment as a whole. This illustrates the interaction between *APOE* ε4 and β-amyloid burden ($P=0.043$).

significant (interaction coefficient = 0.88 (0.76–1.01), $P=0.072$) (Fig. 2b). A similar, but statistically significant, interaction was observed when considering continuous amyloid burden, such that the advantage for ε4 carriers was greater at higher SUVR (interaction coefficient = 0.93 (0.87–1.00) per 0.1 SUVR increment, $P=0.043$) (Fig. 3). Post hoc analyses indicated that this interaction was due to the coefficients for SUVR going in opposite directions within ε4 carriers and noncarriers (that is, lower localization error with increasing SUVR among ε4 carriers versus higher localization error with increasing SUVR among non-ε4 carriers), although neither of these effects was statistically significant (coefficient within ε4 carriers = 0.97 (0.93–1.01) per 0.1 SUVR increment, $P=0.15$; coefficient within non-ε4 carriers = 1.04 (0.99–1.10), $P=0.15$).

The beneficial effect of *APOE* ε4 on localization was consistent across both long and short delays (interaction coefficient = 1.03 (0.97–1.09), $P=0.37$). Additional analyses confirmed that this effect was seen even when considering trials on which the incorrect object was selected (Supplementary Information).

Despite finding sex differences in localization memory (Supplementary Information), and a previous study reporting an interaction between sex and *APOE* ε4 on localization memory[23], we found no evidence for an interaction between sex and *APOE* ε4 (interaction coefficient = 0.99 (0.89–1.09), $P=0.45$).

Hippocampal volume and WMHV did not show statistically significant associations with localization (Table 1).

Misbinding. *APOE* ε4, amyloid status, hippocampal volume, WMHV (Table 1) and amyloid burden (OR = 0.97 (95% CI 0.86–1.09), $P=0.60$) were not associated with misbinding errors. See Supplementary Information for comments on object–location misbinding in relation to previous literature.

Two-dimensional mixture model for sources of localization error. To clarify and extend the results reported above, we used a two-dimensional (2D)-mixture model approach that isolates the contributions of three sources of localization error: misbinding, guessing and imprecision[25,34] (Methods). Its main advantage is the ability to account for random guesses, which can have a potentially large effect on the traditional localization error and misbinding measures. Results for the imprecision parameter in agreement with the traditional localization error metric are as follows:

- ε4 carriers performed better than noncarriers, with significantly lower imprecision (adjusted mean (95% CI): ε4 carriers, 99 pixels (95–103); non-ε4 carriers, 105 (102–108)) (Table 1).
- The reduced imprecision associated with *APOE* ε4 was greater among amyloid-positive than amyloid-negative participants, although this was not statistically significant (interaction coefficient = −11 pixels (−23 to +22), $P=0.090$).
- Imprecision was significantly worse for long delay (adjusted mean (95% CI): long delay, 110 pixels (107–113); short delay, 97 (95–100)) (Table 1).

This confirms that differences in localization error cannot be explained by random guessing, but do indeed reflect differences in the precision of location memory.

Discussion

This study shows that superior performance on a computerized visual working memory task is detectable in *APOE* ε4 carriers at age ~70 years, even in the presence of β-amyloid pathology indicative of preclinical Alzheimer's disease. ε4 carriers had better recall for object identity and recalled locations more precisely, while β-amyloid pathology was independently associated with poorer recall for object identities; our analyses suggest that there is an interaction between *APOE* ε4 and β-amyloid burden for localization. The results support the hypothesis of antagonistic pleiotropy, but also highlight the possibility that the beneficial effects of *APOE* ε4 on specific aspects of cognition may persist into older age[3]. To what extent such a cognitive advantage may explain the survival of the ε4 allele in human populations is an intriguing question.

The superior performance of *APOE* ε4 carriers did not significantly differ according to the length of the delay between encoding and recall (1 or 4 s). This guides us away from attributing the effect to better retention of memory representations over time, instead pointing towards differences in attention and precision of encoding[25]. This is also supported by the patterns of performance we observed on other memory tests in this cohort, with an advantage for ε4 carriers on a verbal memory test with strong attentional and working memory demands, but not on memory tests requiring learning and retention of material over multiple trials (Supplementary Information). This interpretation is consistent with one mechanism proposed for ε4-associated cognitive advantage, which is that ε4 carriers show increased task-related activation in the frontal and parietal regions and correspondingly better performance on tasks requiring attention, short-term memory and top-down cognitive control; such effects have been observed across the life-course[3,11,12,15,16,23,24,35], including in children with Down's syndrome[17]. Therefore, our results may be explained by the attention and frontal/executive demands of this task (with the localization measure being particularly sensitive due to its continuous nature), rather than by visuospatial or memory aspects alone. At older ages, increased frontal activation in ε4 carriers has been proposed to reflect compensatory recruitment since frontal regions are relatively spared from Alzheimer's-disease-related neurodegeneration[3,11,36]. Our finding that the localization advantage for *APOE* ε4 carriers appears relatively greater as amyloid burden increased is consistent with this hypothesis of compensation.

There is currently no consensus on ε4-associated cognitive benefits, and to which domains and functions they may apply[2,3,13,21]. Associations between *APOE* ε4 and poorer cognition are generally

not observed until past middle age[37–39], when they are presumed to reflect the emergence of preclinical pathologies. Additive detrimental effects of *APOE* ε4 and β-amyloid have been observed in older age, with accelerated cognitive decline and disease progression in ε4 carriers[1,40,41], possibly due to additional pathological effects of *APOE* ε4 (for example, on synaptic loss, neuroinflammation and cerebrovascular disease[1,4]). The complex interplay of *APOE* ε4 with different pathologies in the aging brain remains difficult to unravel; our analysis accounted for white matter hyperintensity volume and hippocampal volume (neither of which were associated with task performance), providing evidence that the effects of *APOE* ε4 and β-amyloid on visual working memory are independent of these factors.

One limitation of this study is the relatively small number of trials (24) compared to 100-trial versions used previously[23,24,33,42], although one previous study concluded that a shorter version would be sufficient because group differences were more apparent towards the beginning of the task[42]. Performance on primary outcomes was broadly similar to the longer version, with no floor or ceiling effects, and our results suggest that this short version (duration ~8 min, practical for inclusion in a busy assessment schedule) is sufficient for detection of subtle differences between individuals. Having said that, there are no published data on the test–retest reliability of measures extracted from this task, so future studies examining this point are warranted. One particular strength of our analysis is the availability of life-course data in this cohort, which allowed us to control for factors such as childhood cognitive ability, thus reducing unexplained variability in working memory performance (higher childhood cognitive ability was associated with better recall for object identities and object–location binding (Supplementary Information).

Strengths and limitations relating to the representativeness of Insight 46 participants have previously been discussed[43,44], the main limitations being that all participants were white and the sample was inevitably biased towards those who were willing and able to travel to the research center[45], which may have resulted in under-representation of individuals with cognitive decline or neuropsychiatric symptoms that can be present in preclinical Alzheimer's disease. As previously reported, participants tended to be more highly educated and in better health than their peers not recruited to this substudy, and participants who completed the brain scan were less likely to be obese and to have mental health problems than those with missing neuroimaging data[45]. Because education, obesity and depression are associated with increased dementia risk[46], this raises the possibility that individuals (including *APOE* ε4 carriers) destined to develop Alzheimer's disease or other forms of late-life cognitive impairment may be under-represented in our analyses. The very small number of ε4 homozygotes (consistent with population prevalence) precluded investigation of the dose-dependent effects of *APOE* ε4 (see Supplementary Information for descriptive statistics). Because we conducted a large number of statistical tests with multiple outcome measures, these results require verification in replication studies. Future data collections will include measures of tau pathology, enhancing our ability to draw conclusions about relationships between preclinical Alzheimer's disease and alterations in visual working memory.

In summary, we provide evidence of superior visual working memory in *APOE* ε4 carriers at age ~70 years, even in the presence of subtle cognitive deficits associated with preclinical Alzheimer's disease. This is consistent with the antagonistic pleiotropy hypothesis and suggests that the beneficial effects of *APOE* ε4 on specific cognitive functions may persist into older age.

Methods

Participants in the Insight 46 study—a substudy of the MRC National Survey of Health and Development (NSHD, the British 1946 Birth Cohort)[28]—were assessed at University College London between May 2015 and January 2018. Recruitment procedures, assessment protocols and recruitment flow-charts have been published[28,43,45,47]. In brief, assessments included neuropsychological tests, clinical examination, combined MRI/β-amyloid PET neuroimaging and other biomarker and genetic measures. All assessments were typically completed on one day, although 62 participants had to have their scan rescheduled for a later date (median interval 49 days). The neuropsychological battery comprised standard paper-and-pencil tests and more novel computerized tasks[28,43,44,48], none of which had been administered previously within the NSHD. The study was approved by the Queen Square Research Ethics Committee – London (REC reference no. 14/LO/1173). All participants provided written informed consent.

Stimuli and procedure. The stimuli and procedure of the 'What was where?' task have previously been described in detail[22,33,42,49]. This type of working memory recall precision task has shown convergent validity with traditional measures of working memory span in older adults, and greater sensitivity than these traditional measures to subtle changes in working memory in patients with Parkinson's disease[27]. The participant was seated in front of a 23' DELL Optiplex 9030 all-in-one touch-screen computer. The dimensions of the screen were 1,920 × 1,080 pixels and the approximate distance from the subject's eyes to the center of the screen was 58 cm.

The procedure for the 'What was where?' task is presented in Fig. 1b. In each trial, one or three objects were displayed on the screen at random locations, presented on a black background. Participants were asked to look at the objects and to try to remember their identities and locations. The maximum height and width of the objects was 120 pixels (see Supplementary Information for images of stimuli).

One-object trials are referred to as 'low load' and three-object trials as 'high load', displayed for 1 and 3 s, respectively, to allow time for encoding. This was followed by a blank screen for either a short or long delay (1 or 4 s), and then a test array appeared in which two objects were displayed along the vertical meridian. One of these objects had appeared in the memory array on the previous screen (the target) while the other was a foil/distractor. Participants were instructed to touch the object that they remembered seeing and drag it to the location where they think it was originally presented (Fig. 1b). There was no time limit for reporting the location—the tester pressed the space bar to initiate the next trial when the participant was ready.

Previous studies using the 'What was where?' task have administered at least 100 trials[22–24,33,42,49], but for Insight 46 a shortened version was used containing 24 trials: four low-load and 20 high-load (two low-load with short delay, two low-load with long delay, ten high-load with short delay and ten high-load with long delay). The experiment was preceded by four practice trials—one of each of the load–delay combinations—and the tester ensured that the participant understood the task before continuing.

All objects, including the foils, were drawn from a pool of 60 fractals that were used across the experiment (rendered using http://sprott.physics.wisc.edu/fractals.htm).

The locations of the objects were generated in a pseudorandomized manner by a MATLAB script (MathWorks, Inc.) imposing the following restrictions necessary to allow analysis of localization error, a key outcome of this task: (1) objects were always at least 280 pixels away from each other, to avoid crowding and to ensure that there was a clear zone of 140 pixels around each object (necessary for the calculation of misbinding errors (below)); and (2) objects were at least 200 pixels from the center of the screen and 120 pixels from the edges. The 24 trials were identical for all participants (that is, the same objects were presented in the same locations) but were presented in a random order so that load and delay conditions were interspersed throughout. Using a random order avoids confounding of the results by either practice effects (familiarity with the procedure could result in enhanced performance throughout the task) or interference effects (because objects appear more than once during the task, the foil in the test array could be recognized from a previous trial, which could increase the likelihood of errors in object identification throughout the task).

Outcome variables. *Primary outcomes.* Primary outcomes are illustrated in Fig. 1c. For each trial, an object identification error was recorded if the participant selected the incorrect object from the two-choice array.

Memory for object location was defined in terms of localization error—the distance between the location reported by the participant and the closest of the three original locations from the memory array. This definition takes account of the fact that, in high-load trials, participants may mislocalize the target to the location of a different (unprobed) object from the memory array (that is, they make a misbinding error—see definition below). Previous studies have also calculated gross localization error, which is the distance between the location reported by the participant and the true location of the target in the original memory array. In the case of a misbinding error, the gross localization error could be very large so it is a less pure measure of localization precision. We calculated gross localization error for comparison with previous studies, but it was not used as an outcome measure for statistical analyses.

A misbinding error occurs when a participant correctly identifies the target object but swaps its location with the location of another object. If the target is

positioned within 140 pixels of the location of a different object from the memory array, this is counted as a misbinding error. This threshold was used to ensure that a location could not be attributed to more than one object, because objects were always at least 280 pixels apart. Note that under the low-load condition it is not possible to make a misbinding error because there is only one object in the memory array.

As in previous papers, localization and misbinding errors were analyzed only for trials in which the correct object was identified from the two-choice array[22,33,42,49].

2D-mixture model outcomes. We additionally analyzed performance on the 'What was where?' task using a 2D-mixture model approach that isolates the contributions of three sources of localization error: misbinding, guessing and imprecision[25,34]. In contrast to the traditional localization metric, which considers only the magnitude of localization errors, the 2D-mixture model approach considers two dimensions of error: failure to remember the target location (that is, a misbinding error or a random guess) and imprecision in localization (which applies to both target and misbinding responses). It has shown convergent validity with the traditional metrics and is more accurate at recovering the parameters of simulated data[34]. Because the 2D-mixture model yielded results similar to traditional outcomes, for simplicity we chose to focus this report on the traditional outcomes, with the 2D-mixture model results presented as confirmation that interindividual differences in localization error and misbinding cannot be explained by random guessing. Code for the 2D-mixture model is freely available in the MemToolbox2D package for MATLAB[50]. The model is described in detail[34] but, in brief, a response density equation is defined as follows:

$$P\left(\hat{\theta}\right) = \alpha\psi_\sigma\left(\hat{\theta}-\theta\right) + \beta\frac{1}{m}\Sigma_i^m\psi_\sigma\left(\hat{\theta}-\varphi_i\right) + \gamma\frac{1}{A} \quad (1)$$

θ and ψ are vectors indicating locations on the screen, where $P(\hat{\theta})$ is the probability of finding a response location $\hat{\theta}$, θ is the location of the target, φ_i is the location of the nontarget stimulus i, m is the number of nontarget stimuli, A is the area of the screen and ψ_σ is a bivariate Gaussian distribution with standard deviation σ and zero covariance. The parameters α, β and γ represent the proportion of target responding, misbinding and guessing, respectively. Because α, β and γ must sum to 1, α is not included as a free parameter in the fitting, so the three free parameters are β (misbinding), γ (guessing) and σ (imprecision), estimated using maximum-likelihood methods. Thus, the model isolates and quantifies three different sources of error. Any spatial units can be used, but we used pixels. The model assumes that guesses are uniformly distributed across the entire screen; see Supplementary Information for an exploration of alternatives to this assumption and information about goodness of fit.

To visualize the performance of the model on the raw data, we outputted the probability that each individual response location was (1) target, (2) guess, (3) misbind to the first distractor and (4) misbind to the second distractor. These probabilities were normalized for each trial so that they summed to 1, and then each response was classified into whichever category had the highest probability. The third and fourth categories were then combined, as both represent misbinding. These classifications are illustrated for the complete raw data in Supplementary Information.

Data cleaning. Six participants underwent one trial in which the software did not record whether they had selected the correct or incorrect object, probably caused by the participant touching the screen exactly midway between the two objects. These six trials were excluded.

Life-course and clinical variables. Childhood cognitive ability was measured at age 8 years (or ages 11 or 15 years if earlier data were missing) as a standardized z-score based on tests of verbal and nonverbal ability, as previously described[43].

Educational attainment was represented as the highest qualification achieved by age 26 years, grouped into three categories: no qualification; vocational or O-levels and equivalents; A-levels or degree and equivalents.

Socioeconomic position was derived from participants' own occupation at age 53 years, or earlier if this was missing, coded according to the UK Registrar General's Standard Occupational Classification and classified as manual or nonmanual.

Participants were classified as having a neurological or major psychiatric condition (including dementia and mild cognitive impairment) as previously described[43] (see Fig. 1a for specific diagnoses). Participants not meeting these criteria are herein referred to as cognitively normal and represent a sample free from possible confounding neurological or psychiatric comorbidities. This does not imply that all participants with a neurological or major psychiatric condition necessarily had a measurable cognitive impairment.

Biomarker measures. As previously described[28,47,51], β-amyloid PET and multimodal MRI data were collected simultaneously during a 60-min scanning session on a single Biograph mMR 3 T PET/MRI scanner (Siemens Healthcare), with intravenous injection of 370 MBq of ^{18}F-Florbetapir (Amyvid). β-amyloid deposition was quantified using the SUVR calculated from cortical regions of interest with a reference region of eroded subcortical white matter. A cutoff point for amyloid positivity was determined using a mixture model to define two Gaussians, and taking the 99th percentile of the lower (amyloid-negative) Gaussian at SUVR > 0.6104 (refs. [43,47,51]).

Global WMHV was generated using an automated segmentation algorithm followed by visual quality control[47,52]. Hippocampal volume was generated using the automated segmentation method known as similarity and truth estimation for propagated segmentations, with appropriate manual editing[53]. Total intracranial volume (TIV) was generated using statistical parametric mapping software (SPM12; http://www.fil.ion.ucl.ac.uk/spm)[54].

APOE genotype was determined[28] and classified as either ε4 carrier or non-ε4 carrier. The number of homozygous ε4 carriers (n = 11.3% of the sample) was too small to consider them as a separate group, but descriptive statistics on their performance are provided in Supplementary Information.

Statistical analyses. Analyses were conducted using Stata 15 (StataCorp). Statistical significance was set at the conventional threshold of $P < 0.05$.

To investigate associations between performance on the 'What was where?' task and biomarkers of brain pathologies, analyses included only those participants classified as cognitively normal and for whom complete biomarker data were available (n = 398; Fig. 1a). The rationale for exclusion of participants with neurological or major psychiatric conditions was that these conditions can have varied impacts on cognitive performance that may confound associations between the key predictors of interest (preclinical amyloid pathology and *APOE* ε4) and visual working memory, and the numbers involved were not sufficient to provide power to detect meaningful differences between specific conditions. Multivariable regression models were fitted (details given below), with predictors of amyloid status (positive versus negative), hippocampal volume, WMHV, *APOE* ε4 (carrier versus noncarrier) and delay (short versus long). An additional predictor of load (low versus high) was included for all outcomes except those relating to misbinding, since misbinding cannot occur under the low-load condition. TIV was included as a covariate to adjust for the correlation between brain volume and head size, and we additionally adjusted for sex, age at assessment and the following life-course factors that have previously been shown to predict cognitive performance throughout adulthood in this cohort[43,44,55,56]: childhood cognitive ability, education and socioeconomic position. Adjusting for these factors reduces the unexplained variance in cognitive performance between individuals, which can increase the sensitivity of our analyses to detect subtle effects of APOE ε4 and brain pathologies[43]. Analyses were additionally rerun, replacing dichotomized amyloid status with continuous SUVR. We did not apply a correction for multiple comparisons, following recommendations in the statistical literature[57,58], because this was a hypothesis-driven study motivated by previous literature.

Where between-individual factors were significantly associated with performance, we tested for interactions with delay (short versus long) to investigate whether group differences were due to better retention over time. We also tested for interactions between *APOE* ε4 and amyloid (dichotomous amyloid status and continuous SUVR) to investigate whether the effects of *APOE* ε4 differed according to the burden of amyloid pathology.

Primary outcomes. Analyses were conducted using trial-by-trial data rather than summary scores (for example, mean localization error), to avoid loss of information.

Identification errors (correct versus incorrect) and misbinding errors (yes versus no) were analyzed using generalized estimating equations (GEE) logistic regression models with an independent correlation structure and robust standard errors, to allow for the correlation between repeated measures of the same participant. Results are expressed as odds ratios (OR) for ease of interpretation.

Localization error was analyzed using GEE models, assuming a normal distribution for the dependent variable and an identity link (as with standard linear regression), but including an exchangeable correlation structure and robust standard errors. Localization errors were first log-transformed because the distributions were positively skewed. Model assumptions were tested by examination of residual plots; no departures from assumptions were noted.

2D-mixure model outcomes. To generate the outcome scores for analysis, the mixture model (above) was fitted twice for each participant: once using their responses to short-delay trials (low and high load combined) and once using their responses to long-delay trials (low and high load combined). In regard to the traditional localization metrics, the model included only responses for trials in which the correct object was identified from the two-choice array. This generated a value for the misbinding, guessing and imprecision parameters for each participant under both short- and long-delay conditions. It was not possible to separate responses by load, because the number of low-load trials was too small for reliable estimation of the imprecision parameter (as this is a standard deviation metric). Low-load trials do not influence the estimation of the misbinding parameter, since they do not contain distractors.

The imprecision parameter was analyzed using the same model structure as localization error (above), because it was approximately normally distributed. The guessing and misbinding parameters were analyzed using the same model structure

as identification errors (above), since these represent proportions of responses classified as guesses and misbinds, respectively (see Supplementary Information for examination of goodness of fit and visual illustration of the performance of the model).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All data from NSHD are curated and stored by the Lifelong Health and Aging Unit at UCL. Anonymized data will be shared by request from qualified investigators (https://skylark.ucl.ac.uk/NSHD/doku.php).

Code availability

Code for the 2D-mixture model (MATLAB) is freely available at https://doi.org/10.5281/zenodo.3752705. Code for statistical analyses conducted in Stata is provided in Supplementary Information.

Received: 17 November 2020; Accepted: 17 August 2021;
Published online: 7 October 2021

References

1. Liu, C.-C., Kanekiyo, T., Xu, H. & Bu, G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nat. Rev. Neurol.* **9**, 106–118 (2013).
2. Smith, C. J., Ashford, J. W. & Perfetti, T. A. Putative survival advantages in young apolipoprotein ε4 carriers are associated with increased neural stress. *J. Alzheimers Dis.* **68**, 885–923 (2019).
3. Tuminello, E. R. & Duke Han, S. The apolipoprotein E antagonistic pleiotropy hypothesis: review and recommendations. *Int. J. Alzheimers Dis.* **2011**, 726197 (2011).
4. Safieh, M., Korczyn, A. D. & Michaelson, D. M. ApoE4: an emerging therapeutic target for Alzheimer's disease. *BMC Med.* **17**, 64 (2019).
5. Jack, C. R. et al. Tracking pathophysiological processes in Alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *Lancet Neurol.* **12**, 207–216 (2013).
6. Villemagne, V. L. et al. Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol.* **12**, 357–367 (2013).
7. Baker, J. E. et al. Cognitive impairment and decline in cognitively normal older adults with high amyloid-β: a meta-analysis. *Alzheimers Dement.* **6**, 108–121 (2017).
8. Brookmeyer, R. & Abdalla, N. Estimation of lifetime risks of Alzheimer's disease dementia using biomarkers for preclinical disease. *Alzheimers Dement.* **14**, 981–988 (2018).
9. Byars, S. G. & Voskarides, K. Antagonistic pleiotropy in human disease. *J. Mol. Evol.* **88**, 12–25 (2020).
10. Jasienska, G. et al. Apolipoprotein E (ApoE) polymorphism is related to differences in potential fertility in women: a case of antagonistic pleiotropy? *Proc. R. Soc. B Biol. Sci.* **282**, 20142395 (2015).
11. Duke Han, S. & Bondi, M. W. Revision of the apolipoprotein E compensatory mechanism recruitment hypothesis. *Alzheimers Dement.* **4**, 251–254 (2008).
12. Rusted, J. M. et al. APOE e4 polymorphism in young adults is associated with improved attention and indexed by distinct neural signatures. *Neuroimage* **65**, 364–373 (2013).
13. O'Donoghue, M. C., Murphy, S. E., Zamboni, G., Nobre, A. C. & Mackay, C. E. APOE genotype and cognition in healthy individuals at risk of Alzheimer's disease: a review. *Cortex* **104**, 103–123 (2018).
14. Iacono, D. & Feltis, G. C. Impact of apolipoprotein E gene polymorphism during normal and pathological conditions of the brain across the lifespan. *Aging* **11**, 787–816 (2019).
15. Zink, N., Bensmann, W., Arning, L., Beste, C. & Stock, A. K. Apolipoprotein ε4 is associated with better cognitive control allocation in healthy young adults. *Neuroimage* **185**, 274–285 (2019).
16. Marchant, N. L., King, S. L., Tabet, N. & Rusted, J. M. Positive effects of cholinergic stimulation favor young APOE e4 carriers. *Neuropsychopharmacology* **35**, 1090–1096 (2010).
17. D'Souza, H. et al. Differential associations of apolipoprotein E ε4 genotype with attentional abilities across the life span of individuals with Down syndrome. *JAMA Netw. Open* **3**, e2018221 (2020).
18. Austad, S. N. & Hoffman, J. M. Is antagonistic pleiotropy ubiquitous in aging biology? *Evol. Med. Public Health* **2018**, 287–294 (2018).
19. Abondio, P. et al. The genetic variability of APOE in different human populations and its implications for longevity. *Genes (Basel)* **10**, 222 (2019).
20. Raichlen, D. A. & Alexander, G. E. Exercise, APOE genotype, and the evolution of the human lifespan. *Trends Neurosci.* **37**, 247–255 (2014).
21. Weissberger, G. H., Nation, D. A., Nguyen, C. P., Bondi, M. W. & Han, S. D. Meta-analysis of cognitive ability differences by apolipoprotein e genotype in young humans. *Neurosci. Biobehav. Rev.* **94**, 49–58 (2018).
22. Pertzov, Y., Dong, M. Y., Peich, M.-C. & Husain, M. Forgetting what was where: the fragility of object-location binding. *PLoS ONE* **7**, e48214 (2012).
23. Zokaei, N. et al. Sex and APOE: a memory advantage in male APOE ε4 carriers in midlife. *Cortex* **88**, 98–105 (2017).
24. Zokaei, N. et al. Dissociable effects of the apolipoprotein-E (APOE) gene on short- and long-term memories. *Neurobiol. Aging* **73**, 115–122 (2019).
25. Zokaei, N. et al. Short-term memory advantage for brief durations in human APOE ε4 carriers. *Sci. Rep.* **10**, 9503 (2020).
26. Ma, W. J., Husain, M. & Bays, P. M. Changing concepts of working memory. *Nat. Neurosci.* **17**, 347–356 (2014).
27. Zokaei, N., Burnett Heyes, S., Gorgoraptis, N., Budhdeo, S. & Husain, M. Working memory recall precision is a more sensitive index than span. *J. Neuropsychol.* **9**, 319–329 (2015).
28. Lane, C. A. et al. Study protocol: Insight 46 – a neuroscience sub-study of the MRC National Survey of Health and Development. *BMC Neurol.* **17**, dec2017 (2017).
29. Kuh, D. et al. The MRC National Survey of Health and Development reaches age 70: maintaining participation at older ages in a birth cohort study. *Eur. J. Epidemiol.* **31**, 1135–1147 (2016).
30. Roberts, R. O. et al. Prevalence and outcomes of amyloid positivity among persons without dementia in a longitudinal, population-based setting. *JAMA Neurol.* **75**, 970–979 (2018).
31. Kern, S. et al. Prevalence of preclinical Alzheimer disease: comparison of current classification systems. *Neurology* **90**, e1682–e1691 (2018).
32. Prins, N. D. & Scheltens, P. White matter hyperintensities, cognitive impairment and dementia: an update. *Nat. Rev. Neurol.* **11**, 157–165 (2015).
33. Pertzov, Y., Heider, M., Liang, Y. & Husain, M. Effects of healthy ageing on precision and binding of object location in visual short term memory. *Psychol. Aging* **30**, 26–35 (2015).
34. Grogan, J. P. et al. A new toolbox to distinguish the sources of spatial memory error. *J. Vis.* **20**, 6 (2020).
35. Di Battista, A. M., Heinsinger, N. M. & Rebeck, G. W. Alzheimer's disease genetic risk factor APOE-ε4 also affects normal brain function. *Curr. Alzheimer Res.* **13**, 1200–1207 (2016).
36. Scheller, E. et al. APOE moderates compensatory recruitment of neuronal resources during working memory processing in healthy older adults. *Neurobiol. Aging* **56**, 127–137 (2017).
37. Rawle, M. J. et al. Apolipoprotein-E (Apoe) ε4 and cognitive decline over the adult life course. *Transl. Psychiatry* **8**, 18 (2018).
38. Salvato, G. Does apolipoprotein e genotype influence cognition in middle-aged individuals? *Curr. Opin. Neurol.* **28**, 612–617 (2015).
39. Small, B. J., Rosnick, C. B., Fratiglioni, L. & Bäckman, L. Apolipoprotein E and cognitive performance: a meta-analysis. *Psychol. Aging* **19**, 592–600 (2004).
40. Vermunt, L. et al. Duration of preclinical, prodromal, and dementia stages of Alzheimer's disease in relation to age, sex, and APOE genotype. *Alzheimers Dement.* **15**, 888–898 (2019).
41. Lim, Y. Y. et al. Aβ-related memory decline in APOE ε4 noncarriers: implications for Alzheimer disease. *Neurology* **86**, 1635–1642 (2016).
42. Liang, Y. et al. Visual short-term memory binding deficit in familial Alzheimer's disease. *Cortex* **78**, 150–164 (2016).
43. Lu, K. et al. Cognition at age 70: life course predictors and associations with brain pathologies. *Neurology* **93**, e2144–e2156 (2019).
44. Lu, K. et al. Increased variability in reaction time is associated with amyloid beta pathology at age 70. *Alzheimers Dement. (Amst.)* **12**, e12076 (2020).
45. James, S.-N. et al. Using a birth cohort to study brain health and preclinical dementia: recruitment and participation rates in Insight 46. *BMC Res. Notes* **11**, 885 (2018).
46. Livingston, G. et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* **396**, 413–446 (2020).
47. Lane, C. A. et al. Associations between blood pressure across adulthood and late-life brain structure and pathology in the neuroscience substudy of the 1946 British birth cohort (Insight 46): an epidemiological study. *Lancet Neurol.* **18**, 942–952 (2019).
48. Lu, K. et al. Visuomotor integration deficits are common to both familial and sporadic preclinical Alzheimer's disease. *Brain Commun.* **3**, fcab003 (2021).
49. Pertzov, Y. et al. Binding deficits in memory following medial temporal lobe damage in patients with voltage-gated potassium channel complex antibody-associated limbic encephalitis. *Brain* **136**, 2474–2485 (2013).
50. Grogan, J. P. johnPGrogan/MemToolbox2D: updated release. *Zenodo* https://zenodo.org/record/3752705/export/hx#.YSi5UPlKjIU (2020).
51. Parker, T. D. et al. Hippocampal subfield volumes and pre-clinical Alzheimer's disease in 408 cognitively normal adults born in 1946. *PLoS ONE* **14**, e0224030 (2019).
52. Sudre, C. H. et al. Bayesian model selection for pathological neuroimaging data applied to white matter lesion segmentation. *IEEE Trans. Med. Imaging* **34**, 2079–2102 (2015).

53. Cardoso, J. et al. STEPS: similarity and truth estimation for propagated segmentations and its application to hippocampal segmentation and brain parcelation. *Med. Image Anal.* **17**, 671–684 (2013).
54. Malone, I. B. et al. Accurate automatic estimation of total intracranial volume: a nuisance variable with less nuisance. *Neuroimage* **104**, 366–372 (2015).
55. Richards, M. & Sacker, A. Lifetime antecedents of cognitive reserve. *J. Clin. Exp. Neuropsychol.* **25**, 614–624 (2003).
56. Richards, M. et al. Identifying the lifetime cognitive and socioeconomic antecedents of cognitive state: seven decades of follow-up in a British birth cohort study. *BMJ Open* **9**, e024404 (2019).
57. Armstrong, R. A. When to use the Bonferroni correction. *Ophthalmic Physiol. Opt.* **34**, 502–508 (2014).
58. Althouse, A. D. Adjust for multiple comparisons? It's not that simple. *Ann. Thorac. Surg.* **101**, 1644–1645 (2016).

Acknowledgements

This study was principally funded by grants from Alzheimer's Research UK (nos. ARUK-PG2014-1946 and ARUK-PG2017-1946), the Medical Research Council Dementias Platform UK (no. CSUB19166) and the Selfridges Group Foundation (no. PR/ylr/18575). Genetic analyses were funded by the Brain Research Trust (no. UCC14191). The Florbetapir amyloid tracer was kindly provided by AVID Radiopharmaceuticals (a wholly owned subsidiary of Eli Lilly), who had no part in the design of the study. NSHD is funded by the Medical Research Council (nos. MC_UU_12019/06 and MC_UU_12019/08). The funders of the study had no role in study design, data collection, analysis, interpretation, report writing or the decision to submit the article for publication. T.D.P. was supported by a Wellcome Trust Clinical Research Fellowship (no. 200109/Z/15/Z). A.K. was supported by a Wolfson Clinical Research Fellowship. C.H.S. is supported by an Alzheimer's Society Junior Fellowship (no. AS-JF-17-011). N.C.F. acknowledges support from the UK Dementia Research Institute at University College London, the National Institute for Health Research (Senior Investigator award) and University College London Hospitals Biomedical Research Centre. J.M.S. is supported by University College London Hospitals Biomedical Research Centre, Engineering and Physical Sciences Research Council (no. EP/J020990/1), British Heart Foundation (no. PG/17/90/33415) and EU's Horizon 2020 research and innovation program (no. 666992). We thank participants both for their contributions to Insight 46 and for their commitments to research over the past seven decades. We are grateful to the radiographers and nuclear medicine physicians (A. Groves, J. Bomanji and I. Kayani) at the UCL Institute of Nuclear Medicine, and to the staff at the Leonard Wolfson Experimental Neurology Centre at UCL. We thank D. Marcus and R. Herrick for assistance with XNAT, P. Curran for assistance with data sharing with the MRC Unit for Lifelong Health and Ageing, the DRC trials team for assistance with imaging quality control, M. White for his work on data connectivity and J. Dickson, A. Barnes and D. Thomas for help with imaging.

Author contributions

J.M.S., S.J.C., M.R. and N.C.F. conceptualized and led the Insight 46 study. Y.P. and M.H. designed the visual working memory experiment. K.L., I.M.P. and S.-N.J. collected data for the visual working memory test. T.D.P., C.A.L., A.K., S.E.K. and S.M.B. collected clinical and neuroimaging data. H.M.-S. and A.W. were responsible for study coordination and data management. K.L., S.M.D.H., J.M.S. and S.J.C. conceived the manuscript. J.G. and M.H. developed the 2D-mixture model. K.L. analyzed data and drafted the initial manuscript. J.M.N. provided statistical support. D.M.C., I.B.M., C.H.S. and W.C. generated neuroimaging outcomes. K.L., S.M.D.H., J.G., M.H., J.M.S. and S.J.C. aided in manuscript preparation and interpretation. All authors revised and approved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43587-021-00117-4.

Correspondence and requests for materials should be addressed to Kirsty Lu.

Peer review information *Nature Aging* thanks Duke Han, Miranka Wirth and the other, anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): Kirsty Lu

Last updated by author(s): Aug 10, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The computerised task was run on MATLAB version 2015b using code written by the authors.
Data analysis	Statistical analysis was conducted in Stata 15. Total intracranial volume (TIV) was generated using statistical parametric mapping software (SPM12; http://www.fil.ion.ucl.ac.uk/spm). Code for the 2D-mixture model is freely available in the MemToolbox2D package for MATLAB at https://doi.org/10.5281/zenodo.3752705

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data from the National Survey of Health and Development are curated and stored by the Lifelong Health and Aging Unit at UCL. Anonymized data will be shared by request from qualified investigators (https://skylark.ucl.ac.uk/NSHD/doku.php).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Quantitative cross-sectional. The Insight 46 study is a sub-study of the MRC National Survey of Health and Development (also known as the British 1846 Birth Cohort). 502 participants from the birth cohort were recruited to take part in a full day of assessments including neuroimaging, cognition, clinical function, physical capabilities and blood tests. See full details in our protocol paper (ref 28 in the manuscript).
Research sample	The Insight 46 study is a sub-study of the MRC National Survey of Health and Development (also known as the British 1846 Birth Cohort). Participants were all born in one week in March 1946 and were aged 69-71 at the time of assessment. 51% were female - see more details on demographics in Table 1. The sample is broadly representative of adults of this age born in the UK - details of representativeness have been published (see ref 45 in the manuscript). The rationale for choosing this sample to study relationships between cognition and brain pathologies during the preclinical stages of Alzheimer's disease is: 1) in their early 70s, participants are at an ideal age for studying preclinical Alzheimer's disease since the prevalence of dementia is very low (~3%) but the prevalence of preclinical Alzheimer's brain pathology is around 15-25%; 2) participants are all of almost identical age, which virtually eliminates confounding effects of age differences; 3) participants have a unique resource of data on many aspects of their health, cognition and lifestyle collected prospectively over the last 70 years, which provides the opportunity to account for life-course variables that may explain some of the variance in outcomes of interest in older age. For example, cognitive ability in childhood has consistently been shown to predict cognitive ability across adulthood in the cohort, and we have shown that accounting for this variable increases the sensitivity of our analyses to detect subtle associations between brain pathologies and cognitive performance (see ref 43)
Sampling strategy	Participants were recruited into the cohort in 1946, as a representative sample of babies (n=5362) born across mainland Britain during one week in March 1946. They have been followed up ever since (~2500 remain active in the study). The sample size for the Insight 46 sub-study was set at ~500, based on the constraints of time and funding, and assuming an expected prevalence of amyloid-positivity of ~20%. This is large enough to make meaningful comparisons between the amyloid-positive and -negative groups.
Data collection	The visual working memory task was administered on a computer. The researcher was in the room with the participant during the task; no-one else was present. The researcher was not blind to the study hypothesis, but was blind to participants' amyloid status and APOE genotype at the time of testing.
Timing	May 2015 - January 2018
Data exclusions	As detailed in the manuscript (Figure 1), 486 out of 502 participants completed the visual working memory task (the 16 who did not complete it were excluded from this study). For all analyses presented in the main manuscript, a further 54 participants were excluded as they did not have complete data for the biomarker variables of interest, and a further 34 participants were excluded as they had a neurological or major psychiatric condition which could confound the associations between the predictors and outcome measures. The criteria for these neurological or major psychiatric conditions were agreed within the team to be used consistently across analyses. This resulted in a sample of n=398 for all analyses presented in the main manuscript and for all analyses presented in the Supplementary material (Lu_NatureAging_Supplementary-material_final.pdf), except in section 1. iv) of the Supplementary analyses where we ran statistical models in the full sample of n=486 in order to investigate whether the associations between various demographic and life-course predictors and performance on the visual working memory task materially changed in this larger sample that is more representative of the general population of 70-year-olds. See further details in section 1. iv) of the Supplementary file.
Non-participation	Of the 16 participants who did not complete the visual working memory task, only one declined to do so (the rest were mainly due to technical problems - see Figure 1 for precise details. Of the 54 participants who did not have complete data for the biomarker variables of interest, 25 did declined or withdrew from the brain scan due to claustrophobia (see Figure 1 for other reasons - mainly failures of neuroimaging acquisition / quality control).
Randomization	Participants were not allocated into experimental groups, as this was an observational study (no treatments or interventions were given). All participants completed the same assessments, and key variables of interest were defined such as amyloid status and APOE-e4 genotype. Where participants were grouped by these variables for statistical analysis, we included relevant covariates in the models: sex, age at assessment, childhood cognitive ability, educational attainment, socioeconomic position. As we have described in section 1. i) of the Supplementary file (Lu_NatureAging_Supplementary-material_final.pdf), there were no significant differences in these covariates according to amyloid status or APOE-e4 genotype.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	See above.
Recruitment	Participants were initially recruited into the birth cohort with the consent of their parents, based on being born in the specified week in March 1946. Recruitment into the Insight 46 sub-study in 2015 was carried out by sending invitations to members of the birth cohort who met the basic eligibility criteria (having a defined set of life-course data available, and having previously expressed willingness to come to a London-based clinic visit), then screening by telephone. See full details in ref 45 in the manuscript. All participants provided written informed consent. As detailed in ref 45, participants in the sub-study tended to have higher educational attainment, higher socio-economic position, and better health, compared to those not recruited. As noted in our discussion, this may reflect a self-selection bias where participants in better health, or with more advantaged life circumstances, were more willing or more able to travel to the research centre and undergo the assessments. This is a limitation in the representativeness of our sample, which may have resulted in underrepresentation of individuals with cognitive decline or neuropsychiatric symptoms that can be present in preclinical Alzheimer's disease. This may mean that performance on the visual working memory task was somewhat better than would be seen in the general population. However, the prevalence of amyloid-positivity and APOEe4 in our sample were in line with estimates of population prevalences. As we also noted in our discussion, all participants in the National Survey of Health and Development are white, reflecting the demographics of babies born in the UK in 1946, so they do not represent the ethnic and cultural diversity of the wider population. Therefore, the results may not generalise to more diverse populations; for example, there may be differences by ethnicity in terms of the relationship between APOE-e4 and visual working memory, since the Alzheimer's disease risk conferred by APOE-e4 differs between ethnicities (e.g. see https://alz-journals.onlinelibrary.wiley.com/doi/10.1002/alz.12287)
Ethics oversight	Queen Square Research Ethics Committee - London (REC reference 14/LO/1173)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Magnetic resonance imaging

Experimental design

Design type	Not applicable - no experiments were performed during the scan as it was structural only.
Design specifications	Not applicable - no experiments were performed during the scan as it was structural only.
Behavioral performance measures	Not applicable - no experiments were performed during the scan as it was structural only.

Acquisition

Imaging type(s)	structural
Field strength	3T
Sequence & imaging parameters	3D T1-weighted MPRAGE (magnetization prepared rapid gradient echo) sequence: voxel size = 1.1 mm isotropic, matrix size = 256 × 256 × 208, FOV = 282 × 282, slice coverage = 229 mm, orientation = sagittal, phase encoding direction = A >> P, TE = 2.92 ms, TR = 2000 ms, flip angle = 8 degrees, Acquisition bandwidth = 240 Hz/pix, Water selective excitation pulse, TI = 870 ms. ; 3D FLAIR IR-SPACE (sampling perfection with application-optimized contrasts using different flip angle evolutions) sequence: voxel size = 1.1 mm isotropic, matrix size = 256 × 256 × 176, FOV = 282 × 282, slice coverage = 194 mm,orientation = sagittal, phase encoding direction = A >> P, TE = 402 ms, TR = 5000 ms, flip angle = variable, Acquisition bandwidth = 751 Hz/pix, Water selective excitation pulse, Turbo factor 141, Slice TF 2, T2 sel IR, TI = 1800 ms.
Area of acquisition	Whole brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	TIV: SPM12_r6225 software, brain regions: automatic segmentation with multi-atlas propagation and segmentation (MAPS) with checking and editing by trained individuals, hippocampal: Similarity and Truth Estimation for Propagated Segmentations (STEPS) (Niftypipe commit b8f3272, template manning_hippo_steps_r1.0) with checking and editing by trained individuals
Normalization	Calculating TIV with SPM12 involves pre-aligning to MNI152 with rigid transform
Normalization template	MNI152 (SMP12)
Noise and artifact removal	Gradwarp correction and bias field correction (N4) of structural scans
Volume censoring	Visual quality control of structural MRI by trained raters, scans failed if they contrained too much motion or other artefacts

Statistical modeling & inference

Model type and settings	Not applicable - models were not fitted to the neuroimaging data. Measures of hippocampal volume and white matter hyperintensity volume were used to investigate associations with visual working memory as described in the main manuscript.
Effect(s) tested	Not applicable - models were not fitted to the neuroimaging data. Measures of hippocampal volume and white matter hyperintensity volume were used to investigate associations with visual working memory as described in the main manuscript.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016)	Not applicable - models were not fitted to the neuroimaging data. Measures of hippocampal volume and white matter hyperintensity volume were used to investigate associations with visual working memory as described in the main manuscript.
Correction	Not applicable - models were not fitted to the neuroimaging data. Measures of hippocampal volume and white matter hyperintensity volume were used to investigate associations with visual working memory as described in the main manuscript.

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☒	☐ Multivariate modeling or predictive analysis