Cognitive Reserve Against Alzheimer's Pathology Is Linked to Brain Activity During Memory Formation

Niklas Vockert[1], Judith Machts[1,2], Luca Kleineidam[3,4], Aditya Nemali[1,2], Hartmut Schütze[1,2], Renat Yakupov[1,2], Oliver Peters[5,6], Daria Gref[6], Luisa Sophie Schneider[7], Lukas Preis[6], Josef Priller[5,8,9,10], Eike Jakob Spruth[5,8], Slawek Altenstein[5,8], Anja Schneider[3,4], Klaus Fliessbach[3,4], Jens Wiltfang[11,12,13], Ayda Rostamzadeh[3], Wenzel Glanz[1], Enise I. Incesoy[1,2,14], Stefan Teipel[15,16], Ingo Kilimann[15,16], Doreen Goerss[15,16], Christoph Laske[17,18], Matthias H. Munk[17,19], Annika Spottke[3,20], Nina Roy[3], Michael T. Heneka[3,21], Frederic Brosseron[3], Michael Wagner[3,4], Steffen Wolfsgruber[3,4], Laura Dobisch[1], Peter Dechent[22], Stefan Hetzer[23], Klaus Scheffler[24], Peter Zeidman[25], Yaakov Stern[26], Björn H.Schott[11,12,27], Frank Jessen[3,28,29], Emrah Düzel[1,2], Anne Maass[1,*] , Gabriel Ziegler[1,2,*], and the DELCODE study group

1 German Center for Neurodegenerative Diseases (DZNE), Magdeburg, Germany
2 Institute of Cognitive Neurology and Dementia Research (IKND), Otto-von-Guericke University, Magdeburg, Germany
3 German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany
4 University of Bonn Medical Center, Dept. of Neurodegenerative Disease and Geriatric Psychiatry/Psychiatry, Bonn, Germany
5 German Center for Neurodegenerative Diseases (DZNE), Berlin, Germany
6 Charité – Universitaetsmedizin Berlin, corporate member of Freie Universitaet Berlin and Humboldt-Universitaet zu Berlin, Institute of Psychiatry and Psychotherapy, Berlin, Germany
7 Charité – Universitaetsmedizin Berlin, corporate member of Freie Universitaet Berlin and Humboldt-Universitaet zu Berlin, ECRC Experimental and Clinical Research Center, Berlin, Germany
8 Department of Psychiatry and Psychotherapy, Charité Berlin, Germany
9 School of Medicine, Technical University of Munich; Department of Psychiatry and Psychotherapy, Munich, Germany
10 University of Edinburgh and UK DRI, Edinburgh, UK
11 German Center for Neurodegenerative Diseases (DZNE), Goettingen, Germany
12 Department of Psychiatry and Psychotherapy, University Medical Center Goettingen, University of Goettingen, Goettingen, Germany
13 Neurosciences and Signaling Group, Institute of Biomedicine (iBiMED), Department of Medical Sciences, University of Aveiro, Aveiro, Portugal
14 Department for Psychiatry and Psychotherapy, University Clinic Magdeburg, Magdeburg, Germany
15 German Center for Neurodegenerative Diseases (DZNE), Rostock, Germany
16 Department of Psychosomatic Medicine, Rostock University Medical Center, Rostock, Germany
17 German Center for Neurodegenerative Diseases (DZNE), Tuebingen, Germany
18 Section for Dementia Research, Hertie Institute for Clinical Brain Research and Department of Psychiatry and Psychotherapy, University of Tuebingen, Tuebingen, Germany
19 Department of Psychiatry and Psychotherapy, University of Tuebingen, Tuebingen, Germany
20 Department of Neurology, University of Bonn, Bonn, Germany
21 Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Luxembourg
22 MR-Research in Neurosciences, Department of Cognitive Neurology, Georg-August-University Goettingen, Germany
23 Berlin Center for Advanced Neuroimaging, Charité – Universitaetsmedizin Berlin, Berlin, Germany
24 Department for Biomedical Magnetic Resonance, University of Tuebingen, Tuebingen, Germany
25 Wellcome Centre for Human Neuroimaging, UCL Institute of Neurology, London, UK
26 Cognitive Neuroscience Division, Department of Neurology, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA
27 Leibniz Institute for Neurobiology, Magdeburg, Germany
28 Department of Psychiatry, University of Cologne, Medical Faculty, Germany
29 Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, Koeln, Germany
*authors contributed equally to this work

Abbreviations

ACC anterior cingulate cortex

AD Alzheimer's disease

ADD Alzheimer's disease dementia

ADR AD patient first-degree relatives

CN cognitively normal

CR cognitive reserve

CSF cerebrospinal fluid

DMN default mode network

GM grey matter

MCI mild cognitive impairment

PACC5 Preclinical Alzheimer's Cognitive Composite 5

PC principal component

PCA principal component analysis

PCC posterior cingulate cortex

PFC prefrontal cortex

PL pathological load

SCD subjective cognitive decline

SD standard deviation

TIV total intracranial volume

t-SNE t-distributed stochastic neighbor embedding

Abstract

The cognitive reserve (CR) hypothesis posits that individuals can differ in how their brain function is disrupted by pathology associated with aging and neurodegeneration. Here, we tested this hypothesis in the Alzheimer's disease continuum using longitudinal data from 490 participants of the DELCODE multicentric observational study. Brain function was measured using task fMRI of visual memory encoding. Using a multivariate moderation analysis we identified a CR-related activity pattern underlying successful memory encoding that moderated the detrimental effect of AD pathological load on cognitive performance. CR was mainly represented by a more pronounced expression of the task-active network encompassing the default mode network (DMN), anterior cingulate cortex (ACC) and inferior temporal regions including the fusiform gyrus. We devised personalized fMRI-based CR scores that moderated the impact of AD pathology on cognitive performance and were positively associated with years of education. Furthermore, higher CR scores were associated with slower cognitive decline over time. Our findings suggest maintenance of core cognitive circuits including the DMN and ACC as the primary mechanism of CR. Individual brain activity levels of these areas during memory encoding have prognostic value for future cognitive decline.

1 Introduction

Alzheimer's disease (AD) is biologically characterized by the accumulation of amyloid beta plaques, neurofibrillary tangles made of tau protein, and neurodegeneration[1]. Intriguingly, certain individuals resist clinical progression to dementia in their lifetime despite significant AD pathology in their brains[2]. The cognitive reserve (CR) hypothesis addresses this discrepancy. CR is conceptualized as a mismatch between an individual's brain pathology burden and level of cognitive performance due to cognitive and functional brain mechanisms which are not necessarily accompanied by macroscopic structural brain alterations. Recently, a comprehensive whitepaper presented a unified framework for reserve research and defined CR as the "adaptability (i.e., efficiency, capacity, flexibility) of cognitive processes that helps to explain differential susceptibility of cognitive abilities or day-to-day function to brain aging, pathology, or insult"[3].

In alignment with its definition, the most recent research framework operationalizes CR in the form of a moderator variable[4] (see also `https://reserveandresilience.com/framework/`). As such, it states the requirement of three components for CR research. First, it requires a measure of changes in brain status like brain atrophy or pathology. Second, a quantification of longitudinal changes in cognition theoretically associated with brain status is needed. The third component of this moderation approach is a proposed CR measure, which should moderate the relationship between brain status and cognitive changes. The use of functional neuroimaging methods presents a viable avenue to investigating the neural implementation of CR within this framework. For this purpose, the moderator variable can be represented by the expression of brain activity during cognition using fMRI.

Previous studies have only partially been able to address these aspects, even though a wide range of methodologies has been utilized. Most functional neuroimaging studies on CR have identified regions or networks contributing to CR by correlating their expression (activity/connectivity) with a CR proxy like education or IQ instead of investigating their ability to moderate the relationship between aging- or pathology-related brain changes and cognitive performance. For instance, resting-state functional connectivity profiles have been associated with sociobehavioral CR proxies, manifesting at different levels, including ROI-to-ROI[5], global connectivity of a seed region[6,7], global functional connectivity within a network[8] and employing dimensionality reduction to a ROI-to-ROI connectome[9]. Notably, Van Loenhoud et al.[10] recently employed a task potency method, examining the relationship between whole ROI-to-ROI connectomes in the resting and task state and their association with CR proxies.

Several task-based fMRI investigations pertaining to CR have relied on the Reference Ability Neural Networks Study, wherein participants engaged in 12 cognitive tasks during MRI scanning[10,11,12,13] encompassing three tasks each from the four reference abilities of episodic memory, fluid reasoning, perceptual speed, and vocabulary[14]. This comprehensive approach first facilitated the identification of overlapping regions of brain activity across tasks[11,12,13] and

activity patterns exhibiting correlations with IQ[13] and education[11]. Moreover, Stern et al.[15] observed a distinctive spatial pattern of BOLD activity, the expression of which displayed significant correlations with measures of CR as task load increased.

Among the most notable findings, the default mode network (DMN)[9,10] emerged as a potential CR-related region, alongside its individual components such as the left precuneus[13,16], left posterior cingulate[16,17], precuneus and cingulate[15], and medial frontal gyrus[13]. Left prefrontal cortex activity both within and outside of the frontoparietal network[6] as well as global connectivity of the left frontal cortex[7] were also related to CR. Additionally, there is some evidence for involvement of the anterior cingulate cortex (ACC) in CR[13,17].

However, most previous attempts neglect that a network underpinning CR should be capable of altering the relationship between aging- or pathology-related brain changes and cognitive performance[3]. Moreover, functional neuroimaging studies on CR are more prevalent in aging research, whereas very few investigations have explored CR in the context of neurodegeneration and AD[7,13]. A major challenge in addressing this gap is to obtain brain activity during cognition in large longitudinal cohorts where AD-related pathological burden is thoroughly quantified.

The primary objective of this study was thus to investigate the neural implementation of CR by identifying task fMRI activity patterns associated with cognitive reserve in a large scale multicentric cohort of nearly 500 older individuals along the AD spectrum with the use of the moderation framework. Notably, the cohort was enriched in individuals who still perform normal but are at increased risk for developing AD. To accomplish this, we employed a task fMRI paradigm on memory formation to explore CR within the context of episodic memory. Given that episodic memory is among the earliest and most frequently affected cognitive faculties in dementias like AD dementia[18], memory-related activity patterns hold particular relevance in CR investigations. As the central hub of episodic memory formation and due to its vulnerability in AD, the hippocampus is further of distinct significance for quantification of AD-related neurodegeneration[19]. Our study sought to complement previous approaches by (A) adhering closely to the research framework[4] while (B) identifying a memory activity pattern capable of moderating the impact of AD pathology on cognitive performance. Drawing on insights from prior functional neuroimaging studies on CR, we expected that CR-related activity patterns might encompass regions such as the DMN, frontal regions such as the ACC and task-specific regions like the MTL[20]. Our approach takes advantage of a moderation model in a multivariate fashion (utilizing principal component regression) and effectively condensing the multidimensional AD pathological process (reflecting fluid biomarkers and hippocampal atrophy) into a single pathological load (PL) score. We further derived a neuroimaging-based CR score from an individual's expression of the CR-related fMRI activity patterns and show its alignment with educational attainment, a well-established proxy for CR. Finally, we explored the longitudinal implications of the CR index, meticulously examining its potential to modify cognitive trajectories over time.

2 Results

2.1 Demographics

Our reserve analysis focused on a sample of 490 older participants of cognitively normal (including first degree relatives of AD patients and individuals with subjective cognitive decline) and cognitively impaired individuals (with mild cognitive impairment or AD dementia) who performed task fMRI whose demographics are presented in Table 1 (mean age: 69.7±5.6 years). The sample included slightly more females (53%) than males and was comparably educated (14.6±2.9 years of education). The pathological load (PL) reflected biologically defined AD pathology (Aβ, tau, and hippocampal volume) in one single index ranging from 0 to 1 (see methods for details). The sample mean PL was 0.42±0.3. PL was significantly higher in patients with Alzheimer's disease dementia (ADD) and mild cognitive impairment (MCI) compared to other groups suggesting its validity with respect to clinical diagnosis.

Table 1: **Demographics of the final fMRI sample**. Values represent the mean(sd). *9 participants did not have PACC5 scores. **258 participants did not have PL scores due to missing CSF data. CN = cognitively normal, SCD = subjective cognitive decline, MCI = mild cognitive impairment, ADD = mild Alzheimer's Disease dementia, ADR = AD patient first-degree relatives.

	N	Age [years]	Sex [% female]	Education [years]	PACC5	PL
CN	152	68.89(5.1)	63.2	14.57(2.7)	0.21(0.5)	0.33(0.2)
ADR	51	66(4.7)	56.9	14.49(2.8)	0.14(0.7)	0.3(0.2)
SCD	202	70.01(5.9)	44.6	15.2(2.9)	-0.04(0.7)	0.4(0.3)
MCI	64	72.62(4.8)	53.1	13.44(2.8)	-1.22(0.8)	0.58(0.3)
ADD	21	73.36(5.4)	66.7	13.71(2.8)	-2.97(1.2)	0.84(0.2)
all	490	69.73(5.6)	53.7	14.64(2.9)	-0.19(1.0)*	0.42 (0.3)**

2.2 Pathological load is associated with cognitive performance

The PL score combines CSF measures of amyloid burden and tau pathology with MRI measures of neurodegeneration into a single score. As a robust marker for disease severity along the AD continuum, the PL score exhibited substantial associations with cognitive measures derived from neuropsychological testing. Notably, empirical findings indicated a nonlinear relationship (Fig. 1), prompting a comparison of models with linear and quadratic terms for PL. Both models demonstrated a strong link between PL and PACC5, a composite measure of cognitive performance in the memory domain. The model incorporating a quadratic term displayed a superior fit ($p = 8.78 \cdot 10^{-28}$, standardized regression coefficient β = -0.501, R^2 = 0.420; supplementary information provides comprehensive details).

Since this work focuses on reserve as moderation in terms of interactions with PL, we next tested whether a well-established CR proxy moderated the

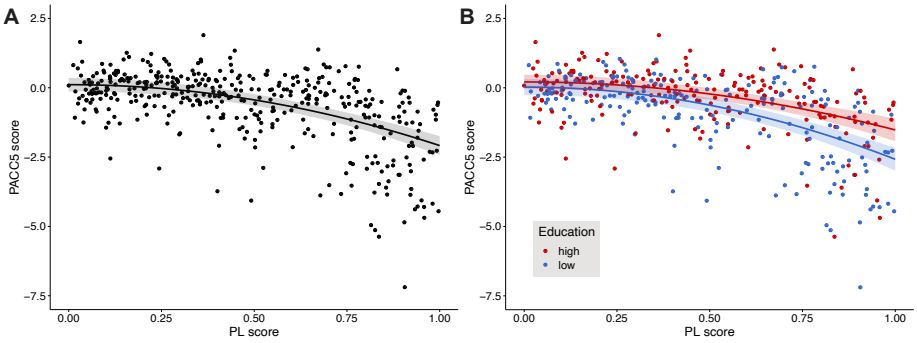

Figure 1: **Cognitive performance declines with higher AD pathological load.** Cognitive performance is represented by the baseline PACC5 score, which was normalized to the unimpaired sample (CN, SCD, ADR). (A) Quadratic model: $PACC5 = b_0 + b_1 \cdot PL^2 + c \cdot COV$. The black line depicts a regression model (with 95% confidence intervals) with a quadratic effect of PL. (B) Same model as in (A), but with an additional interaction term between years of education and the quadratic PL score. Red and blue dots refer to individuals with high and low education, respectively, as obtained by a median split. Red and blue lines are regression models for an individual with average covariate values and 17 (median of the high education group) or 12 years of education (median of the low education group), respectively. Shaded areas refer to the respective 95% confidence intervals.

impact of PL on cognitive performance ($p = 0.0001$, $\beta = 0.752$), which suggested the pivotal role of education in shaping how AD pathology influences cognitive abilities (Fig. 1B). Additionally, in the interaction model, a main effect of the (quadratic) PL score was evident ($p = 2.94 \cdot 10^{-4}$, $\beta = -1.223$), while no further independent main effect of education was observed ($p = 0.575$, $\beta = 0.033$). The moderation model demonstrated an R^2 value of 0.441, further affirming its predictive capability.

2.3 Identification of a CR-related activity pattern

We illustrate brain regions exhibiting heightened activity during encoding for subsequently remembered scenes (Fig. 2A, warm colors) and later forgotten ones (cold colors). In exploring cognitive reserve, we subsequently aimed to identify those spatial patterns (in terms of local voxel-level weights) from this activity contrast that might moderate the impact of a subject's AD pathological load on cognitive performance using a multivariate moderation approach that predicts performance (see methods). Through cross-validation, we determined that the optimal number of principal components (PCs) for the model was 7, yielding a median cross-validation $R^2 = 0.311$ (see Fig. S5).

Our investigation next unveiled patterns of brain regions contributing both positively (depicted in 2B, warm colors) or negatively (cold colors) to the moderation of AD pathology. The former indicates that greater memory encoding activity (more positive or less negative) is linked to superior cognitive performance despite the presence of pathology. Conversely, in regions contributing negatively to the moderation patterns, more negative or less positive activity

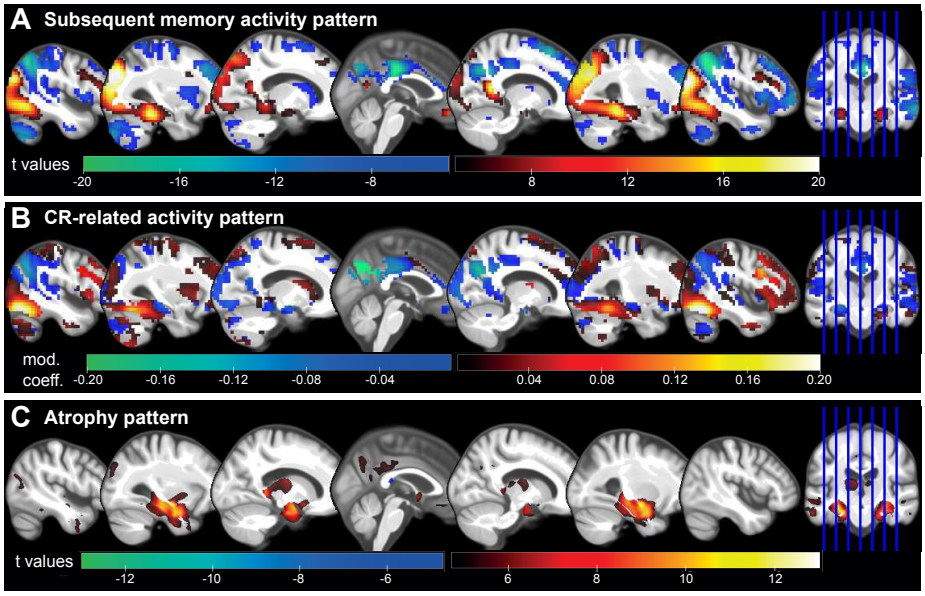

Figure 2: **CR-related activity pattern that moderates effects of pathology**. (A) Activation (hot colors) and deactivation (cold colors) during encoding of subsequently remembered compared to subsequently forgotten stimuli as identified by t-contrasts of the subsequent memory regressor in the whole fMRI sample. T values of voxels with $p_{FWE} < 0.05$ are shown. (B) Group level CR-related activity pattern that when expressed minimizes effects of AD pathology on cognitive performance as identified via a multivariate approach. The net contribution (moderation coefficient; positive/hot and negative/cold colors) of every voxel to the CR-pattern is displayed (unthresholded). (C) Atrophy pattern in the whole baseline DELCODE sample as obtained by a VBM GM analysis of CN vs ADD patients. T values of voxels with $p_{FWE} < 0.05$ are shown.

aligns with better cognitive performance amidst increased pathological burden. In other words, individuals with elevated pathology demonstrated better-than-expected cognitive performance when their BOLD signal differences between subsequently remembered and forgotten stimuli were substantial within regions bearing corresponding colors in Figs. 2A and B. These findings highlight the complex interplay between neural activation patterns and cognitive resilience.

To validate and explore the obtained CR-related activity pattern we then identified clusters with significant contributions (to moderation of pathology) using bootstrapping (Fig. 3 and Tab. 2). Brain regions with the most positive moderation effects were located bilaterally in the inferior temporal and inferior occipital cortices including the fusiform gyri and small parts of right parahippocampal cortex (clusters 1&2). To a weaker extent, parts of the frontal cortex also contributed positively to CR, especially bilateral inferior frontal gyri including opercular, triangular and orbital parts (clusters 3-5) as well parts of right PFC (cluster 6). The strongest negative contribution to moderation were observed in bilateral precuneus, cuneus, posterior cingulate cortex (PCC; cluster 7), bilateral inferior parietal cortex around the angular gyrus (clusters 8&9), and

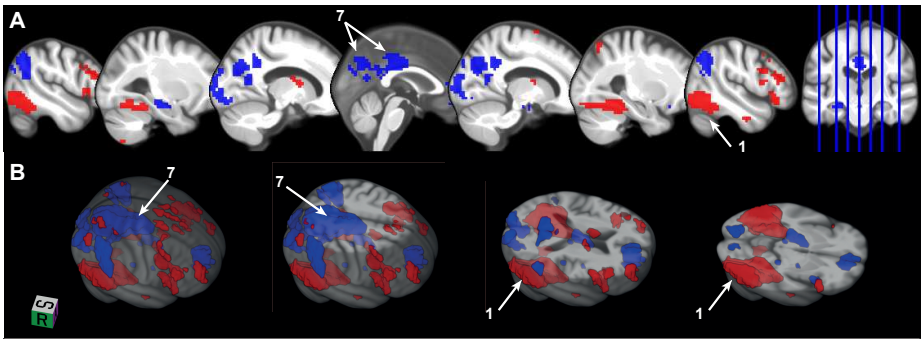

Figure 3: **Significant regions in the CR-related activity pattern**. Several clusters of voxels were determined via bootstrapping whose contribution to the CR pattern (w_i) was found to be significant ($p < 0.05$, see methods). Displayed numbers refer to the clusters described in Tab. 2 with peaks in the following brain structures. 1: right inferior temporal cortex, 7: left precuneus.

primary visual cortex (clusters 10&12). Slightly weaker coefficients were found in the anterior cingulate cortex (cluster 11), the left hippocampus (cluster 13) and medial orbitofrontal cortex (cluster 14).

Interestingly, CR-related activity patterns did not predominantly reflect regions showing atrophy in the DELCODE cohort (mostly found in the hip-

Table 2: **Significant clusters in CR activity pattern**. Structures and peak voxels were identified in MRIcroGL, using the aal (Automated Anatomical Labeling) atlas. w_i refers to the CR coefficient of a voxel i. Type refers to the concordance/discordance between the valence (sign) of the CR coefficient and the subsequent memory contrast coefficient as shown in Fig. 4. For instance, a concordant region is one where a higher (lower) activity reduces effects of pathology and which is typically activated (deactivated) during the task. Clusters smaller than 50 voxels (voxel size: 3.5x3.5x3.5mm) have been omitted.

#	**Mean w_i**	**Size [voxels]**	**Type**	**Peak[x,y,z]**	**Peak Structure**
1	0.107	590	Concordant	49,-63,-17	Temporal_Inf_R
2	0.104	438	Concordant	-49,-66,-17	Occipital_Inf_L
3	0.081	104	Concordant	49, 14, 28	Frontal_Inf_Oper_R
4	0.064	90	Concordant	-49, 25, 24	Frontal_Inf_Tri_L
5	0.057	84	Discordant	49, 25, -7	Frontal_Inf_Orb_R
6	0.060	64	Discordant	35, 59, 7	Frontal_Mid_R
7	-0.106	894	Concordant	0,-60, 32	Precuneus_L
8	-0.072	242	Concordant	-45,-63, 35	Angular_L
9	-0.056	201	Concordant	52,-56, 39	Parietal_Inf_R
10	-0.055	152	Discordant	-7,-91, -3	Calcarine_L
11	-0.053	107	Concordant	4, 49, -3	Frontal_Med_Orb_R
12	-0.057	105	Discordant	14,-91, 3	Calcarine_R
13	-0.059	58	Discordant	-28,-21,-14	Hippocampus_L
14	-0.057	52	Discordant	-3, 56, -7	Frontal_Med_Orb_L

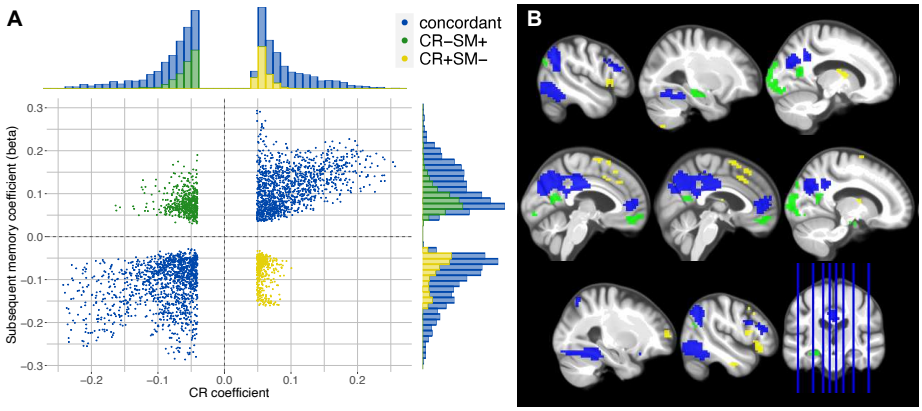

Figure 4: **CR pattern and the subsequent memory contrast predominantly align**. (A) The scatter plot displays the CR coefficients w_i and subsequent memory contrast coefficients (beta) for every voxel with significant contribution to CR. They form three groups: 1. A concordant where both coefficients have the same sign (blue); 2. positive CR coefficient, but negative subsequent memory beta (CR+SM-; yellow); 3. negative CR coefficient, but positive subsequent memory beta (CR-SM+; green). The histograms display the frequency of the voxels in the corresponding groups. The grey dashed line separates the four quadrants. (B) The CR-related activation pattern is shown color-coded corresponding to the colors shown in panel A.

pocampus and medial temporal lobe, Fig. 2C) with minor overlaps in left hippocampus, precuneus and PCC. Overall, this lack of overlap between identified CR-related regions and regions of strongest atrophy was supported by a low correlation of -0.102 (Figs. 2B and C).

To delve deeper into our comprehension of the identified pattern, we conducted an examination of the overlap between the CR-related activity pattern and the generic subsequent memory activity pattern (illustrated in Fig. 2A). A substantial concurrence between these patterns was observed in the most extensive clusters of the CR-related activity pattern, notably in regions such as the precuneus, posterior cingulate cortex (PCC), angular gyrus, ACC and inferior temporal areas (Fig. 4, blue). Within those regions a higher reserve is reflected by an increase in the 'amplitude' of the task-related activation/deactivation. However, it was also apparent that reserve is not uniformly contributing across this task-active network and that certain regions exhibiting significant (de)activation during successful memory encoding did not substantially contribute to cognitive reserve at all (such as portions of the parietal, frontal, temporal, particularly occipital cortex, as well as the cerebellum and basal ganglia).

Moreover, a striking observation emerged in a few brain regions where the valence of the coefficients did not align, a phenomenon we term "discordant" (Tab. 2, Fig. 4). For instance, clusters surrounding the calcarine sulci, bordering the cuneus and precuneus, displayed activation during successful encoding but exhibited a negative contribution to CR (CR-SM+, green in Fig. 4). This

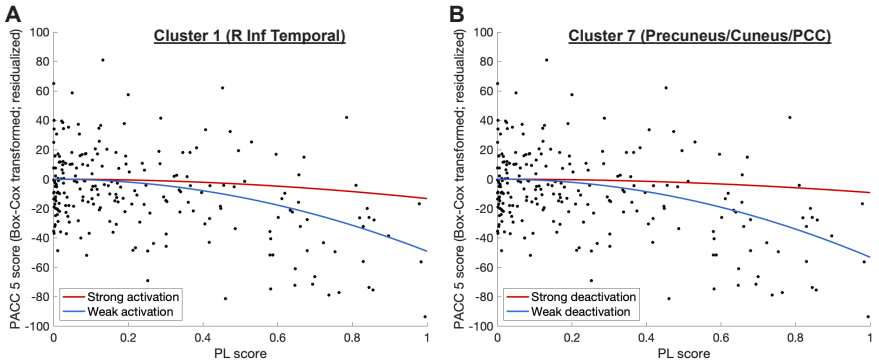

Figure 5: **Subsequent memory-related activity moderates relationship between PL and PACC5**. The relationship between the PL score and the PACC5 score (Box-Cox transformed and residualized for covariates) is moderated depending on the subsequent memory-related activity in two previously identified clusters (see Tab. 2 or Fig. 3). (A) Moderation effect of activation in cluster 1 located around the inferior temporal cortex including fusiform gyrus (positive moderation coefficients). (B) Moderation effect of deactivation in cluster 7 including bilateral cuneus and precuneus as well as posterior cingulate (negative moderation coefficients). The red lines in both panels depict the predicted PACC5 score for an individual with an activation 1 SD above the mean, the blue lines for an activation 1 SD below the mean in the respective cluster. Black dots represent the individual subjects' values for PL and (transformed + residualized) PACC5.

trend was also observed in the left hippocampus and medial orbitofrontal regions. Conversely, positive contributions to CR were evident in certain right frontal areas, such as the insula and mid/superior orbitofrontal cortex (clusters 5&6; CR+SM-, yellow in Fig. 4), despite subsequent memory-related deactivation. Taken together, the correlation between both patterns stood at 0.384, underlining that predominantly more of the typical i.e. activation/deactivation can support cognitive functioning while region-specific multifaceted relationships between these neural signatures and cognitive reserve might exist.

Next, we exemplify how subsequent memory-related activity moderates the detrimental effect of pathology (PL score) on cognitive performance in two brain regions located in right inferior temporal cortex (Fig. 5A) and around bilateral cuneus/precuneus/PCC (Fig. 5B), respectively (taking clusters 1 and 7 from Fig. 3 and Tab. 2). The moderation effect has unveiled a notable phenomenon: as levels of pathological load (PL) rise, the disparities in cognitive performance between individuals with high and low levels of (de)activation become increasingly apparent. Among individuals with high PL, those with high levels of brain activity (i.e. activation in temporal cortex and deactivation in posterior medial cortex) have cognitive ability at the level of individuals with low PL.

2.4 CR score moderates effects of pathology on cognitive performance, also longitudinally

Utilizing the CR-related activity pattern that we identified above, we next derived individualized CR scores. To ascertain its validity as an indicator of cognitive reserve, we expected it to (1) moderate the effect that pathology has on independent cognitive performance measures; (2) moderate longitudinal cognitive decline; and (3) be correlated with sociobehavioral proxies of CR according to the consensus research criteria[3].

Our results affirm the first criterion, demonstrating a moderation effect of the CR score on the relationship between the (quadratic) PL score and cognitive performance across various cognitive tests (Fig. 6A). This moderation effect was evident for the latent memory factor ($p = 8.38 \cdot 10^{-12}$, $\beta = 0.381$), the domain-general factor ($p = 2.15 \cdot 10^{-8}$, $\beta = 0.325$) and the PACC5 score ($p = 9.15 \cdot 10^{-15}$, $\beta = 0.447$), which was originally used in identifying the CR-related activity pattern. Importantly, this moderating influence was observed not only in individuals with cognitive impairment, including MCI and AD patients but also in cognitively unimpaired individuals, emphasizing that the functional activity patterns associated with CR benefit a broad spectrum of cognitive abilities (memory factor: $p = 3.91 \cdot 10^{-5}$, $\beta = 0.273$; domain-general factor: $p = 0.0010$, $\beta = 0.220$; PACC5: $p = 3.77 \cdot 10^{-6}$, $\beta = 0.301$

Furthermore, in an analysis encompassing the remaining sample lacking a PL score (due to missing CSF data), a weaker yet significant moderation effect of the CR score on the association between hippocampal volumes and cognitive

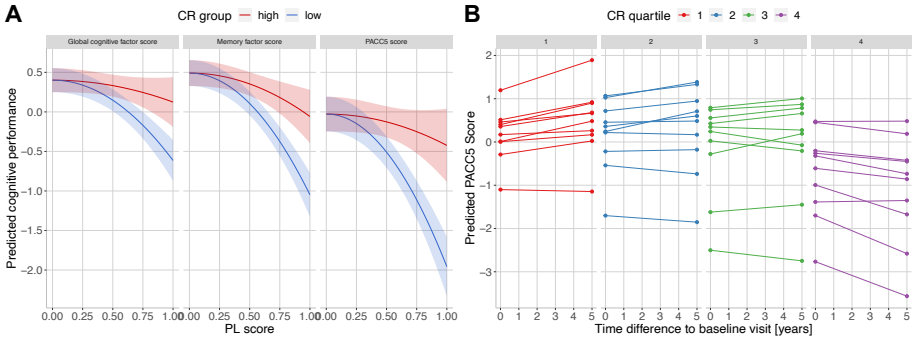

Figure 6: **CR score is linked to cognitive performance cross-sectionally and longitudinally**. (A) The relationship between the PL score and cognitive performance at baseline is moderated by the CR score. Cognitive performance is represented by three different scores: a global cognitive factor score, a memory factor score and the PACC5 score (previously used for identification of the CR-related activity pattern). Cognitive performance was predicted using the respective regression model for an average individual with high (above median; red curve) or low (below median; blue curve) CR score. Shaded areas denote 95% confidence intervals. (B) The longitudinal trajectories of cognitive performance are influenced by the baseline CR score. From each quartile of the CR score, the highest/lowest 10 individuals were selected and their PACC5 scores at a 5 year follow-up were predicted using the previously described LME (see methods section) fitted on the original longitudinal data.

performance was observed in the form of the latent memory factor ($p = 0.018$, $\beta = 0.131$) and PACC5 ($p = 0.011$, $\beta = 0.145$). This moderation effect was not evident for the domain-general factor when considering hippocampal volumes alone instead of PL ($p = 0.147$, $\beta = 0.083$).

In a longitudinal context, the CR score emerged as a significant moderator of rates of change of PACC5 ($p = 0.002$, $\beta = 0.045$), signifying that the CR score, and consequently activity patterns during memory retrieval, hold the potential to influence cognitive trajectories (Fig. 6B). Moreover, the CR score exhibited a positive correlation with education across the entire sample ($p = 0.011$, $r = 0.114$). Collectively, these findings robustly support the contention that the obtained CR score is intimately associated with cognitive reserve, both in cross-sectional and longitudinal assessments.

3 Discussion

In this study, we combined multiple ideas to investigate the neural implementation of cognitive reserve utilizing task fMRI data from a substantial sample comprising 490 participants. First, we employed the most contemporary research criteria governing CR assessment via functional neuroimaging[3,4]. Second, in order to enable this moderation approach we reduced the dimensionality of AD biomarkers in a non-linear fashion, introducing a novel data-driven index quantifying Alzheimer's disease-related pathological load. Third, we pioneered a novel multivariate approach to modeling reserve, which uncovered a task-related functional activity pattern capable of moderating the impact of brain pathology on cognitive performance. Fourth, we provided both cross-sectional and longitudinal validation of the proposed activity pattern of cognitive reserve. Our findings illuminate a compelling connection: older individuals whose brain responses during successful memory encoding more closely align with this identified pattern exhibit diminished cognitive deficits when faced with AD pathology. Moreover, a more pronounced expression of this activity pattern was associated with a slower rate of cognitive decline over longitudinal follow-ups.

In healthy young individuals and older adults, episodic memory encoding is associated with a highly replicated "canonical" pattern of brain activation in some regions and deactivation in other regions[21,22]. We found that a more pronounced expression of this canonical activation/deactivation pattern was associated with higher cognitive reserve. CR was especially characterized by a stronger activation during memory encoding in inferior occipital and inferior temporal areas including the fusiform gyrus, i.e. parts of the ventral visual stream. Some frontal areas showed a similar contribution to CR, though to a smaller extent. CR was further characterized by stronger deactivation in the posterior cingulate cortex, precuneus, cuneus and lateral parietal cortex including angular gyrus, regions that are considered to be part of the DMN[23].

These findings shed new light on the neural implementation of CR. As the majority of brain regions showed concordant activity for CR and successful memory encoding, cognitive reserve primarily seems to be associated with continued recruitment of core cognitive circuits. This indicates that some individuals are able to maintain functional integrity in parts of the core cognitive circuitry despite the presence of AD pathology. Generally, significant decreases in fMRI activity in regions of the DMN have been shown to co-occur with amyloid deposition in older adults[24,25]. Furthermore, AD has been characterized with impairment of regional cerebral blood flow and regional glucose metabolism during resting state, predominantly in temporo-parietal regions[26,27]. Hence, the ability to maintain core fuctional curcuits might represent resilience against pathological changes like neurodegeneration and Aβ accumulation, possibly accompanied by conservation of glucose metabolism. Our findings suggest that this ability is central to CR. The neural mechanisms underlying this ability are still unclear. One possibility is that pathology within core circuitry can be counteracted by non-affected neural populations. This possibility is supported by the pattern of spread of tau pathology within brain tissue where affected and

intact neurons coexist in the same region. Another possibility is that there are individual differences in pathology that we have not quantified. These include inflammation, vascular supply and clearance. It is possible that individuals who are capable of maintaining function in core circuitry despite of tau, amyloid and hippocampal neurodegeneration have less expression of these additional pathologies. Thirdly, genetic mechanisms could contribute. For instance, a human leukocyte antigen locus associated polymorphism was recently reported to provide resilience against tau pathology, indicating a modifying role of adaptive immune responses[28]. In conclusion, rather than relying on the recruitment of additional brain regions as a compensatory mechanism, our findings point towards CR factors operative within core circuits themselves.

Some brain regions of the canonical episodic memory activity network were not associated with CR. Visual areas showed strong activation in the subsequent memory contrast due to the visual nature of the memory task, but did not contribute substantially to cognitive reserve. Likewise, the cerebellum was deactivated during successful encoding, but showed essentially no contribution to CR. These regions have not been discussed much in the context of CR, although there has been scattered evidence for a contribution of inferior and middle occipital regions[12] and the cerebellum[13].

Encoding-related activity in the hippocampus was discordant with cognitive reserve. Thus, although the hippocampus is well-known to be activated during successful memory encoding, weaker left hippocampal activity during encoding was associated with better cognitive performance in the presence of AD pathology. This aligns with observations regarding hyperactivity of the hippocampus in an Aβ- and especially tau-dependent manner that is not related to better cognitive performance[24,29,30]. An absence of pathology-related overactivation in the hippocampus might actually be beneficial for cognitive performance and clinical progression. The calcarine sulci and medial orbitfrontal regions also show discordant acitivity for successful memory encoding versus CR. Some frontal regions including the insula and mid/superior orbitofrontal cortex were deactivated during memory encoding and have positive moderation coefficients, indicating better cognitive performance with weaker deactivation. Yet, it is also conceivable that decreased activity in these regions itself is not actually beneficial for cognitive performance, but systematically co-occurs with beneficial activity changes in other regions.

Generally, the strongest negative contributions to CR were observed in the cuneus, angular gyri, PCC and particularly the precuneus. This is in line with a large body of evidence highlighting the role of the DMN in cognitive reserve. For instance, deactivations of the left precuneus[13,16] and posterior cingulate[16] were associated with CR in previous studies. Moreover, the precuneus together with the cingulate gyrus contributed negatively to some aspects of a CR-related fMRI pattern[15]. Connectivity-based methods provide additional evidence for DMN contribution. For instance, using a task-potency method, which captures a brain region's functional connectivity during task performance after adjusting for its resting state baseline, the DMN has been found to be the predominant contributor to a task-invariant CR network[10]. Additionally, Stern et al.[9]

suggested that connections involving the DMN might be weaker at rest in individuals with higher IQ. Our findings provide further evidence that stronger deactivation of some DMN regions is related to CR.

Albeit weaker in magnitude, the ACC and medial (pre-)frontal cortex, another region attributed to the DMN, showed negative contributions to CR as well. The ACC has also previously been identified in the context of cognitive reserve, e.g. as part of the task-invariant CR pattern of Stern et al.[13]. Moreover, greater volume and metabolism in the ACC were found to be related to higher levels of education[5]. It was further identified as part of a 'resilience signature' whose metabolism was associated with global cognitive performance in cognitively stable individuals over 80 years.[31]

The strongest positive contributions to CR were observed in the fusiform gyri and surrounding temporal to inferior occipital regions. With respect to the fusiform gyrus, there has been both evidence for negative as well as positive contributions to CR.[12,15] Further, some frontal regions have been proposed to play a role in cognitive reserve. For instance, Franzmeier et al.[7] discovered that global connectivity of the left frontal cortex attenuated the relationship of precuneus FDG-PET hypometabolism on lower memory performance in amyloid-positive individuals with MCI. The left frontal cortex also showed positive contributions to CR in our study. Likewise, left prefrontal cortex connectivity both within and outside the frontoparietal network has been found to correlate with fluid intelligence as a proxy of CR[6].

The expression of this CR activity pattern in an individual, as represented by the task-derived CR score, further fulfills the latest research criteria on CR. First, the CR score moderates the effect of pathological load on cognitive performance. Hence, individuals with lower cognitive reserve scores show a stronger nonlinear decline in their cognitive abilities with increasing pathological burden compared to individuals with higher levels of cognitive reserve. The CR score retains its disease-moderating characteristic in the context of multiple different cognitive scores like an independent composite memory measure as well as a very broad measure of cognitive abilities spanning learning and memory, language, visuospatial abilities, executive function and working memory. This reveals a certain robustness of the moderating effect of the CR score and supports its validity. Importantly, the CR score influenced the change of cognitive trajectories over longitudinal follow-ups, stressing its significance not only for present cognitive abilities, but also for their development over time. Furthermore, the CR score was related to education, even though the correlation was found to be rather low to moderate. On the one hand, this could mean that our CR score might capture cognitive reserve incompletely due to the apparent task-dependency. On the other hand, the correlation should not be close to 1 either, since education itself is only a proxy of CR. Thus, education and CR as identified via functional neuroimaging approaches do share parts of their variance, but are also partially independent.

Taken together, the moderating effect of the obtained CR score and its relation to another sociobehavioral CR proxy suggest it as a valid, even though incomplete representation of overall cognitive reserve. It also provides evidence

that the underlying network indeed contributes to CR, at least in context of the incidental encoding task at hand.

This study has a number of shortcomings. The approach was enabled by dimensionality reduction of ATN via the t-SNE method, which provided us with a useful tool for quantifying pathological load. Yet, the PL score is a purely cross-sectional construct that is agnostic for the order of events along the disease progression towards Alzheimer's disease and it may be an oversimplification to represent the ATN system of AD biomarkers by a single variable. Likewise, while hippocampal atrophy is a key feature of AD, it is on oversimplification to represent neurodegeneration solely by hippocampal measures. Strong associations with the three biomarkers (see supplementary) as well as with cognitive performance nevertheless suggest the PL score as a meaningful index of overall disease severity (rather than disease progression) for the purpose of this study. Despite its limitations, t-SNE or other non-linear dimensionality reduction tools like multidimensional scaling or Laplacian Eigenmaps among many others (see e.g. Van Der Maaten et al.[32]) might be useful in many other studies investigating multidimensional disease-related phenomena. Furthermore, our multivariate regression approach relied on a reduction of fMRI data complexity via PCA. As a consequence the moderation analysis might represent an incomplete characterization of CR in the context of successful memory encoding. Moreover, one might consider learning the moderating function of brain activity $f(A)$ and its interactions with pathology (see methods Eq. 1) more directly using neuronal nets and other data-driven approaches[?] . Additionally, CR is treated as a static measure in this study, contrasting with CR's conceptualization as a dynamic entity, susceptible to variation over time[33,34]. However, our approach theoretically allows to represent CR in a dynamic manner by incorporating longitudinal fMRI data. Last, we remark the dependency of the CR-related activity pattern on the task and contrast at hand. We recognize the efforts of task-invariant approaches to identify an underlying pattern of CR that is task-independent and note that the presented multivariate model could be extended to accommodate multi-task data as well. Nonetheless, apart from task-specific components the presented CR activity pattern most likely also contains task-invariant components, e.g. the DMN and ACC. Furthermore, since our contrast probes memory, which is the earliest and most strongly affected faculty in AD, our specific CR-related activity pattern is also of great significance.

In summary, using a multivariate approach to modeling CR we have identified a memory encoding-based activity pattern of cognitive reserve in the context of successful memory encoding according to the latest research definitions. We provide further support for the strong evidence for the hypothesis of a generic role of the DMN and ACC in cognitive reserve. Additionally, we identified regions less commonly associated with cognitive reserve like the fusiform gyrus and some frontal regions. Overall, our findings suggest an enhanced maintenance of core cognitive circuits as the primary mechanism of cognitive reserve. Consequently, interventional efforts should be focused on maintaining the functionality of core cognitive circuitry over the compensatory enhancement of other brain regions with the aim to ameliorate future cognitive decline. Ultimately,

more longitudinal studies are necessary to assess the degree of dynamics of CR over time and its ability to modulate trajectories of cognitive decline. Furthermore, CR patterns have only been assessed on the group-level, assuming CR works similarly across all individuals. Upcoming approaches should account for individual differences in functional (re)organisation, considering individual-level expressions of cognitive reserve.

4 Methods

4.1 Sample

The sample is part of the DZNE-Longitudinal Cognitive Impairment and Dementia Study (DELCODE) study, a multicentric observational study of the German Center for Neurodegenerative Diseases (DZNE). It focuses on the characterization of subjective cognitive decline (SCD) in patients recruited from memory clinics, but additionally enrolled individuals with amnestic MCI, mild AD dementia patients, AD patient first-degree relatives (ADR), and cognitively normal (CN) control subjects. Participants were scheduled for annual follow-up appointments over five years. More detailed information about the study has been provided previously.[35] Our analyses were based solely on the baseline measures of the participants, with the exception of annually acquired cognitive data, which was used to assess cognitive trajectories and how they might be modified depending on an individual's cognitive reserve. The whole baseline sample comprised 1079 participants, of which 558 participants had undergone an MRI session including the fMRI task. 442 participants had cerebrospinal fluid (CSF) data available and could thus be used for creating the PL score. The final fMRI sample used in the subsequent CR analysis consisted of 490 participants after quality control and outlier exclusion, of which 232 had CSF measures and thus a PL score. Of the 490 participants, 152 were CN, 202 had SCD, 64 MCI and 21 had a clinical diagnosis of AD dementia. The sample also contained 51 first-degree relatives of AD patients.

CN was defined as having memory test performances within 1.5 SD of the age-, sex-, and education-adjusted normal performance on all subtests of the CERAD (Consortium to Establish a Registry of AD test battery). ADR had to achieve unimpaired cognitive performance according to the same criteria. SCD was defined as the presence of subjective cognitive decline as expressed to the physician of the memory center and normal cognitive performance as assessed with the CERAD.[36] Participants were classified as MCI when displaying an age-, sex-, and education-adjusted performance below –1.5 SD on the delayed recall trial of the CERAD word-list episodic memory tests. MCI patients were non-demented and had no impairment in daily functioning. Finally, only participants with a clinical diagnosis of mild AD dementia[37] obtaining ≥ 18 points on the Mini Mental State Examination (MMSE) were included in DELCODE. All participants were 60 years or older, fluent speakers of German and had a relative who completed informant questionnaires. Exclusion criteria are described in Jessen et al.[35].

The study protocol was approved by Institutional Review Boards of all participating study centers of the DZNE.[35] The process was led and coordinated by the ethical committee of the medical faculty of the University of Bonn (trial registration number 117/13). All participants provided written informed consent.

4.2 Cognitive tests

An extensive list of all neuropsychological tests administered in DELCODE is provided elsewhere.[35] In our analysis we use composite scores from those tests, namely the Preclinical Alzheimer's Cognitive Composite 5 (PACC5)[38], a neuropsychological composite measure designed to index cognitive changes in the early phase of AD, and a latent memory factor derived from a confirmatory factor analysis (details in Wolfsgruber et al.[39]). The factor analysis yielded five factors for different cognitive domains: learning and memory, language, visuospatial abilities, executive function and working memory. These were further combined to a domain-general global cognitive factor in the form of their mean value. Prior to calculation of the PACC5 scores, the five subitems (see Papp et al.[38] for a specification of the subtests) were z-transformed using the mean and standard deviation of the cognitively unimpaired sample consisting of CN, ADR and SCD participants. Nine of the fMRI-participants lacked PACC5 test scores. One of them also had missing factor scores. In terms of the PACC5 score, for most of the subjects data from multiple time points was available (68, 125, 72, 107, 96, 17 participants with 1, 2, 3, 4, 5, 6 time points, respectively) which was used to model the cognitive trajectories longitudinally.

4.3 CSF measures

Procedures of CSF acquisition, processing, and analysis in the DELCODE cohort have previously been described.[35] Here, we focused on $A\beta_{42:40}$ and phospho-tau181 (p-tau) as CSF measures of amyloid β and tau pathology. Of note, these were used as continuous measures.

4.4 MRI acquisition

MRI data was acquired with Siemens scanners (3 TIM Trio systems, 4 Verio systems, one Skyra and one Prisma system) at 10 different scanning sites. The current analysis was performed using T1-weighted (3D GRAPPA PAT 2, $1mm^3$ isotropic, 256x256 px, 192 slices, sagittal, ca. 5min, TR 2500ms, TE 4.33ms, TI 1100ms, FA 7°), T2-weighted (optimized for medial temporal lobe volumetry, 0.5x0.5x1.5 mm^3, 384x384 px, 64 slices, orthogonal to the hippocampal long axis, ca. 12min, TR 3500ms, TE 353ms) images and a task fMRI protocol (2D EPI, GRAPPA PAT 2, $3.5mm^3$ isotropic, 64x64px, 47 slices, oblique axial/AC-PC aligned, ca. 9 min, TR 2580ms, TE 30ms, FA 80°, 206 volumes). For more details see previous publications.[35,40]. For task fMRI, all sites used the same 3000 MR-compatible LCD screen (Medres Optostim) matched for distance, luminance, color and contrast constant across sites, and the same response buttons (CurrentDesign). All participants underwent vision correction with MR-compatible goggles (MediGlasses, Cambridge Research Systems) according to the same standard operating procedures. SOPs, quality assurance and assessment were provided and supervised by the DZNE imaging network.

Subjects performed a modified version of an incidental visual encoding task

using pictures of indoor and outdoor scenes.[40,41] After familiarization with two "Master" scenes (one indoor, one outdoor) outside of the scanner, participants were presented with 44 repetitions of the Master scenes (22/22) and 88 novel scenes (half outdoor, half indoor) in the MRI scanner using the software Presentation (Neurobehavioral Systems Inc.). Participants were instructed to classify each scene as either indoor or outdoor by pressing a button. Each scene presentation lasted 2500ms, with an optimized inter-trial jitter for statistical efficiency. After a retention delay of 60min, memory was tested outside of the scanner with a 5-point recognition-confidence rating for the 88 former novel scenes and 44 new distractor scenes, to assess successful incidental memory encoding. A response of 1 referred to "I am sure I have not seen this picture before", a 5 meant "I have definitely seen this picture before" and 3 referred to "I don't know".

4.5 Image processing

FreeSurfer 6.0 (`http://surfer.nmr.mgh.harvard.edu/`) was used to obtain measures for hippocampal volumes by combining T1- and T2-weighted images using a multispectral analysis algorithm[42]. Total intracranial volumes (TIV) were obtained using the CAT12 toolbox (version 12.6)[43] in SPM12 r7771 (Wellcome Center for Human Neuroimaging, University College London, UK).

fMRI data preprocessing and analysis were performed using SPM12 and MatlabR2016b/MatlabR2018a. The image preprocessing followed standard procedures: (1) Slice time correction; (2) realignment and unwarping using voxel-displacement maps derived from the fieldmaps; (3) segmentation into grey matter, white matter and CSF; (4) coregistration of functional images to the structural; (5) normalization of the functional images to a population standard space via geodesic shooting nonlinear image registration; (6) normalization to MNI space via an affine transformation; (7) spatial smoothing of the functional images with a 6mm isotropic Gaussian kernel.

In this study, we focused on reserve patterns based on the so called subsequent memory effect, also referred to as successful (memory) encoding , which considers the BOLD-activation during encoding of a stimulus as a function of its subsequent remembering. Following recent methodological research, we decided to model the subsequent memory effect parametrically (see Soch et al.[20]) as opposed to categorically. Higher β values of the subsequent memory contrast images indicate a stronger modulation of the local voxel-based BOLD signal according to the form of the parametric modulator (here arcsine; see below), i.e. a larger difference in BOLD during encoding of later remembered compared to neutral ("unsure") or later forgotten stimuli.

In the first-level general linear model (GLM), all novel scenes were collected into a single onset regressor and a parametric modulator with an arcsine-transformation was applied resulting in the subsequent memory regressor: $\arcsin(\frac{x-3}{2}) \cdot \frac{2}{\pi}$ for a given confidence rating x. A previous study has revealed evidence that this parametric modulator, which puts higher weights on definitely forgotten (1) or remembered (5) items in comparison to probably forgotten (2) or remembered (4) items, is the best choice for a theoretically derived parametric

modulator in the same task-design.[20] The first-level GLM further included the onsets of the Master scenes and covariates including the six motion regressors from the realignment and a CSF-based nuisance regressor. Including nuisance regressors from regions with noise/artifact signal has been shown to increase the sensitivity of BOLD-fMRI studies.[44] In order to obtain a time series for the CSF nuisance regressor, the first eigenvariate of the BOLD time-series was extracted from an anatomical CSF mask. The CSF mask was obtained by thresholding the MNI shoot template of CSF tissue probabilities with a conservative value of 0.9 and eroding it once.

Additional smoothing with a 6mm Gaussian kernel was applied to the subsequent memory contrast images to improve the signal to noise ratio in the heterogeneous large clinical sample. In view of the multivariate setting of our analysis and the required dimensionality reduction of the (high-dimensional) memory contrast images, an inclusion of potential noise components seemed particularly problematic. Hence, we focus on regions with significant subsequent memory contrast activation and deactivation ($p_{FWE} < 0.05$; illustrated in Fig. 2A; 13695 voxels) and therefore excluded regions that might reflect more substantial noise. The obtained "task-active mask" was used to restrict all subsequent fMRI-based reserve analyses.

To enhance the signal-to-noise ratio, we opted for stringent outlier exclusion criteria, predicated on behavioral and task-related fMRI metrics. Individuals were excluded if either of the following was true: (1) Errors in the indoor/outdoor judgement > 8. (2) Absolute response bias > 1.5 in their confidence rating. (3) Framewise displacement (FD) was above 0.5mm in a single EPI or above 0.2mm in more than 2% of the EPIs. (4) An individual had extreme outliers in the β values of more than 10% of the voxels of their (GM-masked) regressor image. 68 individuals were excluded based on these criteria, leaving an fMRI sample with 490 individuals. More comprehensive details regarding this selection process can be found in the supplementary methods.

4.6 One-dimensional pathological load score

We base our multivariate reserve model of CR on a dimensional approach to individual pathological load. More specifically, we here extend the ATN classification system[45] to continuous measures by focusing on joint variation across A, T and N simultaneously. This enables a simplified biological assessment of the individual pathological state, which is likely to be on the continuum from healthy to AD and also avoids difficult a-priori choices for cut-offs. The utilized ATN measures were the following: CSF $A\beta_{42:40}$ ratio (A), CSF p-tau (T) and hippocampal volumes (N). The latter were represented by the sum of their bilateral volumes, divided by the subject's TIV.

All three variables were normalized with their respective means and standard deviations to ensure similar scaling. Due to potential nonlinearity of the disease progression trajectory along the AD continuum in 3D ATN space, a nonlinear dimensionality reduction method called t-distributed stochastic neighbor embedding (t-SNE)[46] was employed in order to reduce the dimension to one,

yielding a single PL score per subject (for more details see supplementary). For this purpose, the scikit-learn library 0.23.2 in Python 3.7 was utilized. The resulting PL score ranged from 0 (minimal AD pathology) to 1 (maximal AD pathology in the sample).

4.7 Multivariate reserve model of brain activity patterns

According to the recent consensus definition, a network that underlies CR should moderate the effect of brain pathologies on cognitive performance[3]. The examination of this moderation effect represents the essence of our multivariate model of CR. More specifically we further study CR in the context of fMRI activity patterns during memory encoding as represented by first level GLM contrast images, one per subject, quantifying their encoding success (for details see section 4.5).

First, if one assumes (scalar) brain activity in a region is given by A and pathological load by PL, then a trivial (linear) moderation model that enables testing a CR effect of activity on cognitive outcomes y could be described as

$$y = b_0 + b_1 \cdot A + f(A) \cdot PL + \epsilon \tag{1}$$

with intercept b_0, main effect of activity b_1 and some function $f(A)$. For CR to improve performance there might be (1) a (linear) additive effect of activity (e.g. $b_1 > 0$ for activations) and/or (2) a (per se non-linear/multiplicative) moderation effect where pathology affects cognition in terms of the slope of PL being a function of activity ($f(A) \neq const$). In principle, it would follow that regional brain activity could (by means of intervention or individual predispositions) be optimized with respect to showing improved performance and/or minimize the detrimental effect of pathology. We are aware that most biological processes are more complex, but for reasons of simplicity we here further focus on the case where the above slope is a simple linear function of activity i.e. $f(A) = b_2 + b_3 \cdot A$ with main effect of pathology b_2 and interaction/moderation effect b_3. Please note that in what follows PL is just used as a quadratic term, as it has been identified as a better predictor of PACC5 compared to a linear term (see section 2.2).

Second, it would be feasable to implement this approach in a mass-univariate (voxel-based) manner that enables testing whether a region (in isolation) contributes to CR in above described ways (1) and/or (2). However, since the subsequent memory contrast activity is representing spatially correlated patterns in many brain areas reflecting distributed information processing we opt for a multivariate approach (also avoiding multiple testing and increasing sensitivity). We therefore further assume the above activity A that might contribute to CR (via $f(A)$) to be reflected by patterns of voxelwise subsequent memory contrast images (in task-active areas), i.e. $b_3 \cdot A = \sum w_i \beta_i$ with linear (group-level) weights w_i describing a voxel's potential contribution to CR and its contrast value β_i. Please note that we assumed free weight parameters to be positive or negative enabling potentially enhanced and reduced activations

serving reserve processes. This approach generalizes above ideas of univariate CR as well as brain-based multivariate cognition-prediction models by asking if there is any activity pattern (which a subject could more or less express) that facilitates CR by means of a moderation of pathology effects.

Third, due to the large number of parameters (w_i) we implement the multivariate reserve model by means of representing the subsequent memory contrast images by projections on P-order principal components basis functions (images) obtained from PCA. This resembles an application of principle components regression with principal components being used for quantification of patterns of (1) main effects as well as (2) the moderation effect representing CR in a narrower sense. The finally applied multivariate reserve model is a prediction model of cognitive performance including main effects of activity patterns and their interactions with pathology:

$$y = b_0 + \sum_{p=1}^{P} b_{1,p} \cdot PC_p + (b_2 + \sum_{p=1}^{P} b_{3,p} \cdot PC_p) \cdot PL^2 + c \cdot COV + \epsilon \qquad (2)$$

with PACC5 cognitive performance scores y, individual pathological load score PL, component scores PC_p for corresponding (PCA) eigenimages p and COV representing the covariates age at baseline, sex, TIV and a binary dummy variable indicating MRI acquisition at a specific site. Since PL scores were dependent on the availability of CSF measures, the model was restricted to a subsample of 232 participants (see Fig. S2). Age and TIV were mean-centered. Education was deliberately not chosen as a covariate in this context due to its role as a CR proxy. PACC5 scores had been transformed with a Box-Cox transformation (lambda = 2.8) in order to achieve a closer approximation of the model's residuals to the normal distribution. The coefficients $b_{3,p}$ represent the moderation effect indicative of CR according to the consensus framework[3]. The optimal number of principal components P required to characterize reserve patterns based on subsequent memory contrast images is a free hyperparameter in the multivariate reserve model. It was optimized using a 10-fold cross-validation approach described in the supplementary. In the next step, PCA was performed on the complete (mean-centered) functional data using the optimized value of P. The multivariate reserve model (Eq. 2) with the previously identified optimal number of principal components was estimated on the whole data set, obtaining coefficients for each principal component.

Then, the approach enables us to project the obtained moderation coefficients $b_{3,p}$ for the PCs back into the image space for the purpose of illustration and to determine the net moderation effect $w_i = \sum_{p=1}^{P} b_{3,p} V_{p,i}$ of all voxels i with eigen-images V_p obtained from PCA. Therefore, w_i represents how (strong) the local subsequent memory contrast in voxel i (i.e. the activity associated with successful memory encoding) contributes to moderation of the effect that pathology has on cognitive performance differences and thus its potential role for cognitive reserve. Finally, we introduce a useful reserve score (denoted as CR_{score}) as the amount of how an individual's successful encoding activity aligns with the reserve pattern we identified on the group level by aggregating indi-

vidual contrast images using the reserve weights over all voxels in the mask: $CR_{score} = \sum w_i \beta_i$.

4.8 Statistical Analyses

4.8.1 PL score

For validation purposes, the association between the retrieved PL scores and PACC5 cognitive test scores was examined, including education and the covariates age at baseline, sex and site of data acquisition. Both models with PL as a linear and a higher-order quadratic predictor of PACC5 were tested and their performance was compared in terms of their explained variance (R^2 value). The quadratic predictor was tested due to visual indications for a quadratic relationship between PL and PACC5. Furthermore, such quadratic relationships have been observed in similar contexts, for example between age and brain structure (see e.g. Ziegler et al.[47]). Instead of testing a full quadratic model including a linear term, we restricted ourselves to finding a single predictor of disease severity in order to avoid a further increase in the complexity of the subsequent multivariate reserve model and thus aid its interpretability. Furthermore, inclusion of an additional linear term did not provide substantial increases in explained variance ($\Delta R^2 = 0.014$; in comparison: $\Delta R^2 = 0.040$ between the quadratic-only and linear-only model). An additional model assessed an interaction of the PL score with education as a CR proxy according to the assumption that cognitive reserve moderates the effect that brain pathology (PL) has on cognitive outcomes (PACC5).

4.8.2 Multivariate reserve model - voxel-wise inference

Inference on the voxel-level in the context of the multivariate moderation analysis was performed using a bootstrapping procedure similar to the one previously presented in the context of multivariate mediation analysis of Chen et al.[48]. The following steps were done in 5000 iterations of bootstrapping:

1. Create a bootstrap sample of equal size as the original sample used in the multivariate moderation model by randomly resampling subjects with replacement from it.

2. Estimate bootstrap coefficients $\hat{b}_{3,p}$ from Eq. 2 for the bootstrap sample.

3. Obtain individual voxel bootstrap moderation coefficients $\hat{w}_i$.

The coefficients $\hat{w}_j$ were then stacked into a 5000*13695 matrix. Subsequently, the bootstrap distribution of all voxels was sorted by their median values and voxels whose median lies within the second or third quartile were selected. 10 % of all coefficients $\hat{w}_j$ from the selected voxels were randomly sampled and combined to form a pseudo-null distribution. Finally, the pseudo-null distribution was utilized to fit a normal distribution, which can be used to obtain a p value for every voxel based on its coefficient w_i. Voxels with p values below 0.05 were

regarded as significant. Please note that non-identifiability of the sign of the coefficients is, unlike in the multivariate mediation approach, not a problem in the multivariate moderation approach and thus the inference step here deviates slightly from the one in Chen et al.[48].

4.8.3 Validation of the CR score

In order to ensure the validity of the CR score (section 4.7), its moderating effect between the PL score and cognitive performance was tested in a separate moderation model using different cognitive scores on top of PACC5.

$$Performance = b_0 + b_1 \cdot BAE + b_2 \cdot PL^2 + b_3 \cdot CR_{score} \cdot PL^2 + c \cdot COV + \epsilon \quad (3)$$

where CR_{score} represents the subject level weighted sum of moderation coefficients ($b_{3,p}$). Moreover, BAE reflects the simpler additive effect of brain activity on performance, i.e. a score calculated analogously to the CR score but aggregating the additive components $b_{1,p}$ from Eq. 2 instead. In addition to PACC5 we here used a memory factor and a global cognitive factor score as dependent variables to demonstrate that the main result obtained from learning reserve patterns based on PACC5 generalizes to other cognitive scores. This validation analysis was not possible in the subset of participants without a PL score (due to biomarker unavailablility). We instead performed a similar analysis for these participants in which the PL score (in Eq. 3) was replaced solely by neurodegeneration, represented by TIV-corrected bilateral hippocampal volumes (squared). Additionally, the correlation between the CR score and education as a typical CR proxy was assessed across the whole sample. Since above model training and analyses were cross-sectional, as a final validation step, we utilized linear mixed-effects modelling (package lme4 in R) to test longitudinally if the CR score would also predict individual differences of cognitive trajectories in PACC5 (i.e. not explicitly accounting for individual AD pathology). The model included a subject-specific intercept and slope, as a model comparison had suggested a model with both random intercept and slope as superior compared to one with a random intercept alone. The model examined the interaction effect between the CR score and time between measurements (continuous variable). Age at baseline, sex, site of data acquisition and diagnostic group were included as covariates.

Acknowledgement

We thank the members of the DELCODE group in Magdeburg for providing helpful feedback and suggestions with respect to the conducted analyses.

Funding

The study was funded by the German Center for Neurodegenerative Diseases (Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE)), reference

number BN012.

Competing interests

Dr. Stern consults for Eisai, Lilly, and Arcadia. Columbia University licenses the Dependence Scale, and in accordance with university policy, Dr. Stern is entitled to royalties through this license. Dr. Schott is involved in clinical studies by Roche, Biogen, and Hummingbird Diagnostics, but does not receive personal funds from any of them.

References

[1] Jack, C. R. *et al.* NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's and Dementia* **14**, 535–562 (2018).

[2] Katzman, R. *et al.* Clinical, pathological, and neurochemical changes in dementia: A subgroup with preserved mental status and numerous neocortical plaques. *Annals of Neurology* **23**, 138–144 (1988).

[3] Stern, Y. *et al.* Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimer's & Dementia* **16**, 1305–1311 (2020).

[4] Stern, Y. *et al.* A framework for concepts of reserve and resilience in aging. *Neurobiology of Aging* (2022).

[5] Arenaza-Urquijo, E. M. *et al.* Relationships between years of education and gray matter volume, metabolism and functional connectivity in healthy elders. *NeuroImage* **83**, 450–457 (2013).

[6] Cole, M. W., Yarkoni, T., Repovš, G., Anticevic, A. & Braver, T. S. Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *Journal of Neuroscience* **32**, 8988–8999 (2012).

[7] Franzmeier, N., Duering, M., Weiner, M., Dichgans, M. & Ewers, M. Left frontal cortex connectivity underlies cognitive reserve in prodromal Alzheimer disease. *Neurology* **88**, 1054–1061 (2017).

[8] Franzmeier, N. *et al.* Resting-state global functional connectivity as a biomarker of cognitive reserve in mild cognitive impairment. *Brain Imaging and Behavior* **11**, 368–382 (2017).

[9] Stern, Y., Varangis, E. & Habeck, C. A framework for identification of a resting-bold connectome associated with cognitive reserve. *NeuroImage* **232**, 117875 (2021).

[10] van Loenhoud, A. C., Habeck, C., van der Flier, W. M., Ossenkoppele, R. & Stern, Y. Identifying a task-invariant cognitive reserve network using task potency. *NeuroImage* **210** (2020).

[11] Habeck, C. *et al.* The Reference Ability Neural Network Study: Life-time stability of reference-ability neural networks derived from task maps of young adults. *NeuroImage* **125**, 693–704 (2016).

[12] Habeck, C., Eich, T., Razlighi, R., Gazes, Y. & Stern, Y. Reference ability neural networks and behavioral performance across the adult life span. *NeuroImage* **172**, 51–63 (2018).

[13] Stern, Y., Gazes, Y., Razlighi, Q., Steffener, J. & Habeck, C. A task-invariant cognitive reserve network. *NeuroImage* **178**, 36–45 (2018).

[14] Stern, Y. *et al.* The Reference Ability Neural Network Study: Motivation, design, and initial feasibility analyses. *NeuroImage* **103**, 139–151 (2014).

[15] Stern, Y. *et al.* A Common Neural Network for Cognitive Reserve in Verbal and Object Working Memory in Young but not Old. *Cerebral Cortex* **18**, 959–967 (2008).

[16] Bosch, B. *et al.* Cognitive reserve modulates task-induced activations and deactivations in healthy elders, amnestic mild cognitive impairment and mild Alzheimer's disease. *Cortex* **46**, 451–461 (2010).

[17] Stern, Y. Brain Networks Associated with Cognitive Reserve in Healthy Young and Old Adults. *Cerebral Cortex* **15**, 394–402 (2005).

[18] Morris, R. G. & Kopelman, M. D. The Memory Deficits in Alzheimer-type Dementia: A Review. *The Quarterly Journal of Experimental Psychology Section A* **38**, 575–602 (1986).

[19] Small, S. A., Schobel, S. A., Buxton, R. B., Witter, M. P. & Barnes, C. A. A pathophysiological framework of hippocampal dysfunction in ageing and disease (2011).

[20] Soch, J. *et al.* Bayesian model selection favors parametric over categorical fMRI subsequent memory models in young and older adults. *NeuroImage* **230** (2021).

[21] Kim, H. Neural activity that predicts subsequent memory and forgetting: A meta-analysis of 74 fMRI studies. *NeuroImage* **54**, 2446–2461 (2011).

[22] Soch, J. *et al.* A comprehensive score reflecting memory-related fMRI activations and deactivations as potential biomarker for neurocognitive aging. *Human Brain Mapping* **42**, 4478–4496 (2021).

[23] Andrews-Hanna, J. R., Smallwood, J. & Spreng, R. N. The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Annals of the New York Academy of Sciences* **1316**, 29–52 (2014).

[24] Sperling, R. A. *et al.* Amyloid Deposition Is Associated with Impaired Default Network Function in Older Persons without Dementia. *Neuron* **63**, 178–188 (2009).

[25] Buckner, R. L. *et al.* Molecular, Structural, and Functional Characterization of Alzheimer's Disease: Evidence for a Relationship between Default Activity, Amyloid, and Memory. *Journal of Neuroscience* **25**, 7709–7717 (2005).

[26] Frackowiak, R. S. *et al.* Regional cerebral oxygen supply and utilization in dementia. A clinical and physiological study with oxygen-15 and positron tomography. *Brain : a Journal of Neurology* **104**, 753–778 (1981).

[27] Silverman, D. H. *et al.* Positron Emission Tomography in Evaluation of Dementia: Regional Brain Metabolism and Long-term Outcome. *JAMA* **286**, 2120–2127 (2001).

[28] Luo, G. *et al.* Protective Association of HLA-DRB1*04 Subtypes in Neurodegenerative Diseases Implicates Acetylated Tau PHF6 Sequences. *Neurology* **99**, S17–S17 (2022).

[29] Maass, A. *et al.* Alzheimer's pathology targets distinct memory networks in the ageing brain. *Brain* **142**, 2492–2509 (2019).

[30] Huijbers, X. W. *et al.* Tau accumulation in clinically normal older adults is associated with hippocampal hyperactivity. *Journal of Neuroscience* **39**, 548–556 (2019).

[31] Arenaza-Urquijo, E. M. *et al.* The metabolic brain signature of cognitive resilience in the 80+: beyond Alzheimer pathologies. *Brain* **142**, 1134–1147 (2019).

[32] Van Der Maaten, L. J. P., Postma, E. O. & Van Den Herik, H. J. Dimensionality Reduction: A Comparative Review (2009).

[33] Bettcher, B. M. *et al.* Dynamic change of cognitive reserve: associations with changes in brain, cognition, and diagnosis. *Neurobiology of Aging* **83**, 95–104 (2019).

[34] Lenehan, M. E. *et al.* Sending Your Grandparents to University Increases Cognitive Reserve: The Tasmanian Healthy Brain Project. *Neuropsychology* **30**, 525–531 (2016).

[35] Jessen, F. *et al.* Design and first baseline data of the DZNE multicenter observational study on predementia Alzheimer's disease (DELCODE). *Alzheimer's Research and Therapy* **10** (2018).

[36] Jessen, F. *et al.* A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimer's and Dementia* **10**, 844–852 (2014).

[37] McKhann, G. M. *et al.* The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's and Dementia* **7**, 263–269 (2011).

[38] Papp, K. V., Rentz, D. M., Orlovsky, I., Sperling, R. A. & Mormino, E. C. Optimizing the preclinical Alzheimer's cognitive composite with semantic processing: The PACC5. *Alzheimer's & Dementia: Translational Research & Clinical Interventions* **3**, 668–677 (2017).

[39] Wolfsgruber, S. *et al.* Minor neuropsychological deficits in patients with subjective cognitive decline. *Neurology* **95**, e1134–e1143 (2020).

[40] Düzel, E. *et al.* CSF total tau levels are associated with hippocampal novelty irrespective of hippocampal volume. *Alzheimer's and Dementia: Diagnosis, Assessment and Disease Monitoring* **10**, 782–790 (2018).

[41] Düzel, E., Schütze, H., Yonelinas, A. P. & Heinze, H.-J. Functional phenotyping of successful aging in long-term memory: Preserved performance in the absence of neural compensation. *Hippocampus* **21**, 803–814 (2010).

[42] Iglesias, J. E. *et al.* A computational atlas of the hippocampal formation using ex vivo, ultra-high resolution MRI: Application to adaptive segmentation of in vivo MRI. *NeuroImage* **115**, 117–137 (2015).

[43] Gaser, C. & Dahnke, R. CAT - A Computational Anatomy Toolbox for the Analysis of Structural MRI Data (2016).

[44] Behzadi, Y., Restom, K., Liau, J. & Liu, T. T. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage* **37**, 90–101 (2007).

[45] Jack, C. R. *et al.* A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. *Neurology* **87**, 539–547 (2016).

[46] Van Der Maaten, L. & Hinton, G. Visualizing Data using t-SNE. Tech. Rep. (2008).

[47] Ziegler, G., Dahnke, R. & Gaser, C. Models of the aging brain structure and individual decline. *Frontiers in Neuroinformatics* **6**, 3 (2012).

[48] Chén, O. Y. *et al.* High-dimensional multivariate mediation with application to neuroimaging data. *Biostatistics* **19**, 121–136 (2018). `1511.09354`.

[49] Jack, C. R. *et al.* Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *The Lancet Neurology* **9**, 119–128 (2010).

[50] Vogel, J. W. *et al.* Four distinct trajectories of tau deposition identified in Alzheimer's disease. *Nature medicine* 1–11 (2021).