

The effects of bilingualism on hippocampal volume in ageing bilinguals

Article

Published Version

Creative Commons: Attribution 4.0 (CC-BY)

Open access

Voits, T., Robson, H., Rothman, J. and Pliatsikas, C. ORCID: <https://orcid.org/0000-0001-7093-1773> (2022) The effects of bilingualism on hippocampal volume in ageing bilinguals. *Brain Structure and Function*, 227. pp. 979-994. ISSN 1863-2661 doi: 10.1007/s00429-021-02436-z Available at <https://centaur.reading.ac.uk/101557/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1007/s00429-021-02436-z>

Publisher: Springer

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online



The effects of bilingualism on hippocampal volume in ageing bilinguals

Toms Voits¹ · Holly Robson² · Jason Rothman^{1,3} · Christos Pliatsikas^{3,4}

Received: 23 December 2020 / Accepted: 28 November 2021
© The Author(s) 2022

Abstract

Long-term management of more than one language has been argued to contribute to changes in brain and cognition. This has been particularly well documented in older age, where bilingualism has been linked to protective effects against neurocognitive decline. Since memory difficulties are key aspects of this decline, herein we examine potential effects of bilingualism on the hippocampus, a brain structure related to memory that is particularly vulnerable to cognitive ageing. Hippocampal volume has been shown to increase as a result of second language learning and use in younger adults. However, it is unknown if this is maintained throughout the lifespan. We examine hippocampal volume and episodic memory performance in a participant sample consisting of healthy older individuals with a wide range of experiences in exposure and using a second language. Results reveal greater hippocampal volume calibrated to degree of quantified dual language use. Our results mirror those of immersive active bilingualism in younger populations, suggesting that long-term active bilingualism leads to neuroprotective effects in the hippocampus. We discuss this in the context of literature proposing bilingualism-induced brain reserve in the older age.

Keywords Bilingualism · Ageing · Memory · Hippocampus · Structural MRI · Experience-based neuroplasticity

Introduction

Bilingualism has been shown to be a lifestyle enrichment factor correlating with greater neural plasticity across the lifespan, at least under conditions of active and sustained engagement (see Pliatsikas 2020 for review). Directly or indirectly, these outcomes are hypothesised to be a consequence of increased demands for executive and language

control needed to manage more than one linguistic system in a single mind/brain (e.g., Grundy et al. 2017). Moreover, research has shown that ageing bilinguals outperform monolinguals in various domains of executive functioning, such as mental set shifting, updating, and inhibition (e.g., Bialystok et al. 2004; Sullivan et al. 2016, although see Lehtonen et al. 2018). With regard to neuroanatomy in older age, bilingualism has been associated with greater grey matter volume and white matter integrity across brain structures involved in bilingual language control, language learning, and language processing (Anderson et al. 2018a, b; Duncan et al. 2018; Gold et al. 2013). Such findings are of particular importance, as older age is the period in life where cognition and the brain naturally decline.

Typical cognitive ageing is most clearly identifiable in anatomical changes such as reductions in grey matter (GM) volume and/or white matter (WM) integrity, especially in the prefrontal cortex and the hippocampus, and/or decreased neural efficiency (i.e., increased recruitment of implicated networks) in task performance (e.g., Rönnlund et al. 2005; Persson et al. 2006; Giorgio et al. 2010; Nyberg et al. 2010; Bettio et al. 2017; Farokhian et al. 2017). However, there is a general variability in cognitive ageing trajectories across

✉ Toms Voits
toms.voits@uit.no

¹ Department of Language and Culture, UiT the Arctic University of Norway, Hansine Hansens veg 18, 9019 Tromsø, Norway

² Department of Psychology and Language Sciences, University College London, Chandler House, 2 Wakefield Street, London WC1N 1PF, UK

³ Centro de Investigación Nebrija en Cognición, Universidad Nebrija, Calle de Sta. Cruz de Marcenado, 27, 28015 Madrid, Spain

⁴ School of Psychology and Clinical Language Sciences, University of Reading, Harry Pitt Building, Earley Gate, Whiteknights Road, Reading RG6 6AL, UK

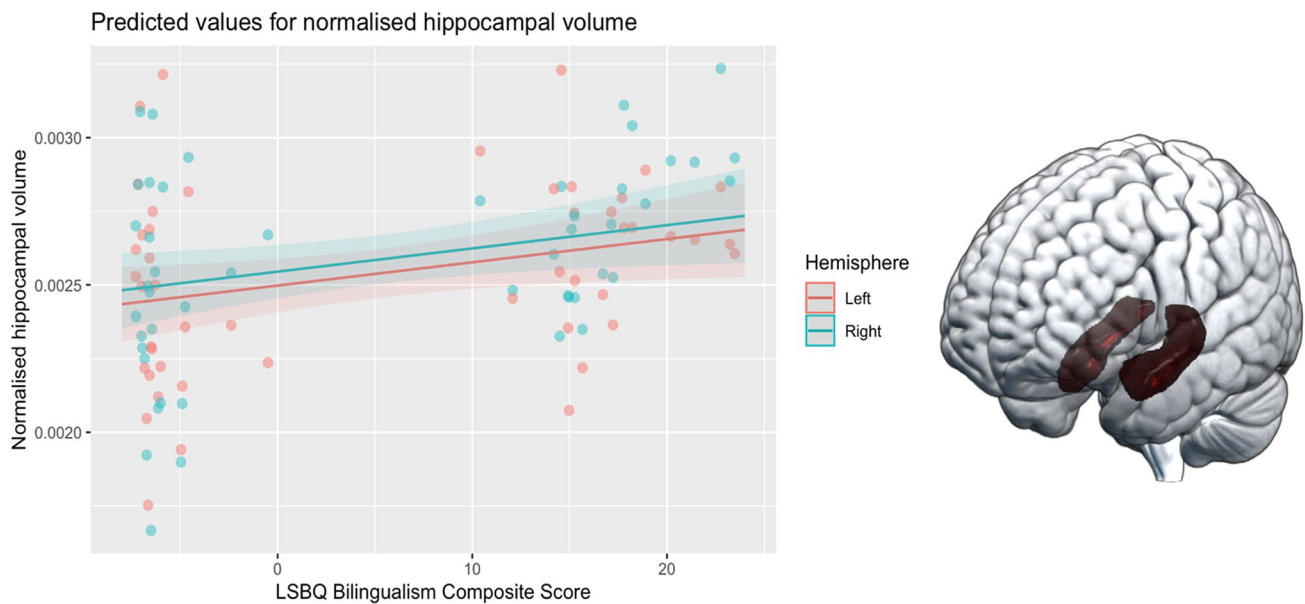


Fig. 1 Plot of the predicted values of normalised hippocampal volumes against LSBQ BCS (Model 3), superimposed on observed data points. See Table 3 for statistical comparisons. Bilateral hippocampus shown on the right-hand side on an MNI template for illustrative purposes

the population (see Cabeza et al. 2018 for review). Some individuals seem to be more resilient to age-related cognitive decline. In addition to genetic factors as determinants of individual differences, variability can be explained by the widely used concepts of cognitive and brain reserve (Stern et al. 2018). Cognitive reserve refers to preserved cognitive ability in the face of neural damage or atrophy, manifesting as better-than-expected cognition in cases of progressive neurodegeneration (Stern 2002). Brain reserve refers to the build-up of neural tissue, as a structural reinforcement of the brain, via volumetric increases caused by neurogenesis or dendritic branching (Valenzuela and Sachdev 2006). In individuals with increased brain reserve, neural decline may take longer before any cognitive and behavioural symptoms manifest.

Brain reserve is typically observed in healthy individuals and linked to a variety of lifestyle enrichment factors, such as higher education, physical exercise, demanding leisure activities, and high occupational attainment (Cabeza et al. 2018; Darwish et al. 2018; Foubert-Samier et al. 2012; Hötting and Röder 2013; Pernecky et al. 2019; Ritchie et al. 2019; Yaffe et al. 2009). Bilingualism also stands out as a potential lifestyle factor for reserve accrual. This is so because the mechanisms implicated in the mental stimulation/exercise required to efficiently maintain, manage, and use multiple languages overlap with those believed to be at the core of accrual for other lifestyle-enrichment factors. Indeed, while some studies report null results (Mukadam et al. 2017; Yeung et al. 2014; Zahodne et al. 2014), an increasing number of studies provide evidence of

bilingualism contributing to the delay of dementia symptom onset in neurodegenerative diseases, most commonly in Alzheimer's disease or Mild Cognitive Impairment (MCI) (e.g., Alladi et al. 2013; Bialystok et al. 2007; Calabria et al. 2020; see Anderson et al. 2020 for critical review).

Active bilingualism has been demonstrated to have implications for episodic memory performance in elderly adults (mean age 80+) (Schroeder & Marian 2012). However, despite compelling reasons to the contrary, few studies have examined the hippocampus, a core element of the episodic memory network, in bilinguals at any age and none specifically in older bilinguals (DeLuca et al. 2019b; Mårtensson et al. 2012; Li et al. 2017). Since (active) bilingualism often correlates to later diagnosis of dementia (see Anderson et al. 2020 for critical review) and diagnosable symptoms often relate to real world memory issues, investigating brain areas potentially underlying this observation—structures and networks where memory is core—is timely and important.

The hippocampus is a bilateral grey matter structure in the medial temporal lobe (see Fig. 1), associated with supporting episodic memory function, but it also underlies other important aspects of cognition, such as recognition, spatial processing, language learning, emotional behaviour, vocabulary acquisition, and mental imagery (Anand and Dhikav 2012; Bellmund et al. 2018; Bird and Burgess 2008; Breitenstein et al. 2005; Ullman 2004). Previous work has linked reductions in hippocampal size with verbal and non-verbal episodic memory performance decline (Gorbach et al. 2017; O'Shea et al. 2016). The hippocampal anatomy is subject to decline in healthy ageing by annual volumetric loss of

0.79–2%, surpassing that of other brain structures (in comparison, annual gross brain volume reduces by 0.2–0.5%) and becomes increasingly rapid with older age (Fjell et al. 2009; Fraser et al. 2015). Moreover, hippocampal atrophy is an established indicator for conversion from healthy ageing to development of mild cognitive impairment (Fotuhi et al. 2012) and from latter to Alzheimer's disease (Apostolova et al. 2006).

Notwithstanding, the hippocampus has been shown to be plastic in response to changes in behaviour. For example, behavioural or physical interventions can impact hippocampal volume and improve memory performance in the older age, effectively reversing age-related hippocampal tissue loss (Erickson et al. 2011; Firth et al. 2018; Lövdén et al. 2012). Early life intellectual enrichment has been linked to increased hippocampus volume (Sumowski et al. 2016), which in turn has been shown to contribute to cognitive resilience in the pathologically ageing brain (Erten-Lyons et al. 2009). In sum, not only does the hippocampus appear to be a malleable brain structure, but its malleability seems to have correlates with behaviour.

Not unlike other areas of the brain that bilingualism is argued to impact, data on how bilingualism affects the hippocampus are somewhat mixed, with most of the literature indicating hippocampal anatomy to be sensitive to bilingual language experience. Mårtensson et al. (2012) examined Swedish interpreter students and found a significant volumetric increase in the right hippocampus following an intensive 3-month language course. Bellander et al. (2016) also reported expansion of the right hippocampus in young Swedish speakers as they acquired Italian vocabulary over the course of 4 months. Interestingly, expansions were not related to the amount of vocabulary acquired, but to the amount of time spent studying the second language (L2), i.e., engagement with additional language learning overall. In a longitudinal study, DeLuca et al. (2019b) tested bilinguals living in an immersive L2 environment for 3 years, and reported significant reshaping of the right hippocampus in the form of simultaneous expansions and contractions of different subfields of the structure. Li et al. (2017) compared hippocampal volumes between highly proficient bimodal Mandarin Chinese–Chinese Sign Language bilinguals and Mandarin Chinese monolinguals (aged 29–67). They showed enlarged hippocampus for the former group, who reported active use of both their languages on a regular basis. In juxtaposition to the above, Olsen and colleagues (2015) did not find any volumetric differences in the hippocampus between 70-year-old bilinguals and monolinguals, although they did find differences in other parts of the temporal lobe.

One reason for some inconsistency in the findings might relate to how bilingualism itself is operationalised across studies [a more general issue in the cognitive neuroscience of bilingualism literature, see Leivada et al. (2021)

for discussion]. Indeed, most studies looking at the effects of bilingualism on neurocognition have treated bilingualism as a discrete, binary variable (whereby one is categorised as monolingual or bilingual). Such an approach fails to acknowledge, much less capture, the dynamic nature of bilingualism and the ensuing potential variability across bilinguals. Recently, there has been a push to unpack individual differences *across* bilinguals. In this manner, bilingualism is treated in a more nuanced way, a *continuum*, by finding ways to qualify and quantify an individual's bilingual experiences (Bak 2016; Bialystok 2017; De Cat et al. 2018; DeLuca et al. 2019a, b, 2020; Gullifer et al. 2018; Luk and Bialystok 2013). Beyond addressing the obvious question of whether or not bilingualism can result in neurocognitive adaptations observable in older age per se, one wants (if not needs) to isolate and better understand the conditions of language exposure/engagement that differentiate individual bilinguals along the trajectory of outcomes (Grundy 2020; Leivada et al. 2021).

The present study

In the context of the above discussion, the present study focuses on individual differences of bilingual experience on the hippocampus in healthy ageing. Highly proficient speakers of English as a second language in long-term immersion and native English-speakers with ranging from limited to no working knowledge of other languages (functional monolinguals) underwent a behavioural and MRI testing battery assessing their memory and hippocampal structure. This was accompanied by a collection of detailed language background information permitting quantification of bilingual experience on an individual level along a spectrum. We treat language experience and regress it as a dynamic, continuous variable within a collapsed group of all participants (functional monolinguals and bilinguals) and then only within the self-identifying bilingual sample, as in Pliatsikas et al. (2021). In this way and in line with calls in the recent literature (Luk and Bialystok 2013; DeLuca et al. 2020; de Bruin 2019), we sidestep two potential comparative fallacies: (a) the assumption that monolinguals and bilinguals form a priori distinct groups and (b) that the members of either group are so similar to one another in relevant sense that that individual variation is trivial. In the case (a) or (b) is true or happens to apply to our sample, the collapsing we performed in our models that run proxies for *language engagement* would reveal this anyway.

We propose and test two hypotheses regarding bilingualism, hippocampal structure, and memory performance. In line with the previous findings in younger populations, we expected to see greater hippocampal volume predicted by increased bilingual exposure/engagement. Second, if language exposure/engagement contributes to increased

Table 1 Demographic information

	Self-reported monolinguals (<i>n</i> = 25)	Self-reported bilinguals (<i>N</i> = 23)	Statistical comparison
Mean age (SD)	65.6 (10.7)	58.5 (6.77)	$p = 0.008136$
Sex	14 F; 11 M	16 F; 7 M	Chi-sq = 0.45078; $p = 0.502$
Education	19.1 (3.59)	20.5 (3.16)	$p = 0.1569$
LSBQ bilingualism composite score	− 5.94 (1.58)	17.0 (3.43)	$p < 0.001$
LSBQ L2 Home score	− 12.3 (2.50)	5.23 (4.52)	$p < 0.001$
LSBQ L2 Social score	− 6.06 (2.86)	51.8 (8.40)	$p < 0.001$
L2 English speaking proficiency (out of 10)	–	8.48 (1.28)	–
L2 English reading proficiency (out of 10)	–	8.95 (0.996)	–
L2 English writing proficiency (out of 10)	–	8.47 (1.37)	–
L2 English understanding proficiency (out of 10)	–	8.82 (1.16)	–
English age of acquisition (years; bilinguals only)	–	10.7 (6.12)	–
Length of L2 immersion (bilinguals only)	–	29.5 (17.2)	–

hippocampal volume, we further predict better episodic memory performance correlated to increased hippocampal volume.

Methods

Participants

Forty-eight healthy older adults (30 females, *mAge*: 62.19, *SD*: 9.62, range 48–84) were recruited for the study. Of these, 23 self-identified as bilingual or multilingual speakers of L2 English (16 females, *mAge*: 58.48, *SD* = 6.77, range 49–73) (henceforth referred to as ‘bilinguals’) and 25 were functionally monolingual native English speakers, some of which had had some experience with other languages (e.g., second language learning at school age), but reported being able to hold a conversation only in English (*mAge* = 65.60, *SD* = 10.68, range 48–84) (henceforth referred to as ‘monolinguals’). All participants were right-handed and reported no neurological disorders or history of speech and language impairments, and they were all resident in the UK at the time of testing. Prior to participation, subjects provided written informed consent and reported no counterindications to MRI scanning. All participants scored within the normal range of the ACE-III (Hsieh et al. 2013), suggesting no indications of cognitive impairment. Behavioural testing and MRI scanning sessions were mostly conducted on the same day, although in some cases, where it was not feasible to conduct all aspects of testing in 1 day, participants returned for a second round of testing at a later date. The maximum time period between the testing sessions was 3 months.

The bilingual participants spoke a variety of first languages but converged on English being an additional language. Most of these participants (*N* = 22) reported speaking

an additional language or languages to English and their respective L1. While this means that these individuals were not strictly bilinguals but brought different language backgrounds and experiences to the table, they all converged on the fact that they have a long-term engagement with bilingualism, while living immersed in an environment where their first language is not a majority one. This means that their bilingual language control processes are actively used, leading to potential changes in neurocognition. In terms of language proficiency, two bilingual individuals reported English to be their most proficient language, 16 reported English as their second most proficient language, three individuals reported English as their third most proficient language, and one reported English as their fourth most proficient language. These participants usually acquired English at school age (*mAoA*: 10.65; *SD*: 6.12; range 0–30). The majority of this group were born outside the UK and had moved to the UK at various ages. Two participants in this group were born in the UK, but did not speak English at home and started learning English upon commencement of formal education. One participant was born in the Netherlands and reported growing up in a bilingual Dutch/English household. Participants in this group had been immersed in their additional language environment for an extended period of time (mean length of residence in the UK = 29.52 years; *SD* = 17.20; range 1–60), and were using English for everyday communication and were competent and highly proficient users of this language (see Table 1). Of the self-reported monolinguals, 13 participants reported some exposure to an additional language, usually at school age. However, none of the monolinguals reported continuous engagement with their additional languages at the present day, mostly advising ‘occasional use while on holiday’. Active engagement with their L2 was normally in a classroom setting during adolescence, decades prior to testing.

The Language and Social Background Questionnaire (LSBQ; Anderson et al. 2018a, b) (see Sect. [Language and Social Background Questionnaire \(LSBQ\)](#) for details) offers bilingualism composite scores under -3.12 as firmly bilingual and scores over 1.22 as firmly monolingual. Individuals scoring between these values lie in a ‘grey area’, with ambiguous language background. Although self-identifying as “monolinguals”, two participants scored between -3.12 and 1.22 in LSBQ; one of these participants had no working knowledge of any other languages but reported growing up in an environment where they were surrounded by other languages. The other participant reported extensive experience with French, although they were not actively engaging in use of French in their everyday life. This variation in scores alone shows the need to move past group comparisons and treat bilingualism as a more nuanced variable. No participants were excluded from analysis based on their linguistic background. For full language and demographic information, split by self-reported ‘*lingualism*’, see Table 1.

Behavioural data collection

Language and Social Background Questionnaire (LSBQ)

The participants completed the language and social background questionnaire (LSBQ) (Anderson et al. 2018a, b). The LSBQ is a questionnaire that allows one to collect detailed information about one’s social (professional attainment, country of birth, etc.) and linguistic background (spoken languages, self-rated proficiency, age, and context of acquisition), and the extent of language use across different contexts (see Mann and de Bruin (2021), for recent work testing and highlighting the effectiveness of the LSBQ). Bilingual experience is quantified via a bilingualism composite score (BCS) as a sum of various quantitative experience-based factors such as extent of L2 use in home and social settings. The BCS allows for measurement and treatment of bilingualism as a continuous variable, as opposed to the more commonly used and now questioned stratification of participants in monolingual and bilingual language groups (see de Bruin 2019; Pliatsikas et al. 2020; Surrain and Luk 2019; Leivada et al. 2021).

The participants completed a paper copy of the LSBQ on their own, but an examiner was present to answer any questions participants may have had and provide clarification as needed. As LSBQ presumes English to be the native or first language by default, the calculations using the factor score calculator were canonical for the native English speakers, whereas the calculations for those with other first languages were altered for their native language to be treated as the baseline, and English regarded as L2. This required inversion of some scores from the questionnaire upon input in the factor score calculator (as in DeLuca, et al. 2019a, b).

NIH toolbox

A modified cognition battery of the NIH Toolbox (NIH-TB; Weintraub et al. 2013) was used to assess the cognitive functioning of the study participants. The NIH-TB is an iPad-based testing battery. For the present study, two tests were of particular interest: the NIH-TB Picture Sequence Memory Test (testing episodic memory performance) and NIH-TB List Sorting Test (testing working memory performance), discussed in detail below. Hippocampus is typically associated with episodic memory performance, while working memory relies on frontal and parietal networks (Nee and D’Esposito 2015). Nonetheless, hippocampal volume has been shown to correlate with performance in the specific NIH toolbox working memory task in ageing populations (O’Shea et al. 2016). Inclusion of two tasks tapping into different memory domains allows us to test for the specificity of the results to episodic memory function with the hippocampus and for involvement of this structure in working memory processes. In the List Sorting Working Memory Test, participants were presented with cartoon pictures of different foods and animals, with accompanying audio presentation and written text naming the item. The participants were then asked to repeat them back to the examiner listing them in size order from the smallest to the biggest. In the first condition, participants were asked to recall stimuli from one category. In the second condition, participants were presented with stimuli from two categories (foods and animals) in mixed order and required to recall the items in size order for each category separately. The number of items in each trial increases until two trials of the same length are failed. All items were of high frequency, easily recognisable, and unambiguous. The test is scored as the total items correct across all trials.

In the Picture Sequence Memory Test, sequences of pictured objects and activities were presented in a particular order. The participants were then asked to reproduce the same order on the screen. The pictures are presented in two trials: one with a 15-step sequence and the other with an 18-step sequence. The second sequence is a repetition of the same 15 items, with three novel items added in the middle of the sequence. The score is derived by the cumulative number of adjacent pairs remembered correctly over the learning trials.

Both NIH toolbox tests were automatically scored with uncorrected standard scores measuring behavioural performance. Moreover, age and education measures, also collected as part of the behavioural data via the NIH toolbox, were included in the analysis as covariates. The education scoring of the NIH toolbox takes into account the highest level of education achieved and estimates years of formal education from it.

Table 2 Neurocognitive measures and outcomes

	Self-reported mono-linguals (<i>n</i> =25)	Self-reported bilinguals (<i>N</i> =23)	Statistical comparison <i>P</i>
Mean ACE-III total score	94.0 (4.67)	94.5 (3.72)	0.6713
Mean ACE-III memory domain score	23.8 (2.57)	25.0 (1.36)	0.04888
NIH-TB Episodic memory score	96.2 (12.4)	103 (11.7)	0.07583
NIH-TB Working memory score	97.5 (10.5)	99.7 (6.99)	0.4097
Total normalised hippocampal volume ($\times 10^3$)	4.90 (0.669)	5.36 (0.445)	0.006894
Left normalised hippocampal volume ($\times 10^3$)	2.44 (0.350)	2.64 (0.254)	0.02545
Right normalised hippocampal volume ($\times 10^3$)	2.46 (0.373)	2.72 (0.246)	0.006478

The significance level for bold should be defined at the level of $p < 0.05$

Addenbrooke's Cognitive Examination (ACE-III)

Participants were asked to complete the Addenbrooke's Cognitive Examination (ACE-III) testing battery (Hsieh et al. 2013). ACE-III is a widely used screening tool for cognitive deficits in Alzheimer's Disease and Frontotemporal Dementia. It is scored out of 100 and covers five cognitive domains—attention, memory, fluency, language, and visuospatial processing. An overall score of less than 82 is suggestive of potential dementia. The domain of primary interest in this study was memory. The tasks tapping into memory are scattered throughout the exam and tap into working, episodic, and semantic memory. More specifically, the participants are asked to recall previously repeated words, memorise and recall a fictional name and address, and recall well-known historically significant people (Bruno and Vignaga 2019). The memory domain is scored out of 26. The score provides a baseline information of one's composite memory performance and was used in addition to the NIH toolbox cognitive battery episodic and working memory tasks. All participants performed within normal limits, indicating typical ageing (see Table 2).

MRI data acquisition

High-resolution T1 anatomical scans were acquired using an MPRAGE sequence on a 3 T Siemens MAGNETOM Prisma_fit MRI scanner, with a 32-channel Head Matrix coil and Syngo software (256 sagittal slices, 0.7 mm slice thickness, in-plane resolution 250×250 , acquisition matrix of 246×256 mm, 224 mm FoV, TR = 2400 ms, TE = 2.41 ms, inversion time = 1140 ms, flip angle = 8°). The scan lasted approximately 10 min.

MRI data processing

Pre-processing

Structural neuroimaging data were pre-processed and analysed with software pipelines in FSL. All T1-weighted scans

were then anatomically pre-processed using the fsl_anat pipeline in FSL 5.0.9 (Smith et al. 2004). This involves a standard use of various MRI processing tools including the brain extraction tool (BET) used for skull stripping the raw T1 images and bias field correction as part of the pipeline. Bias field-corrected T1 images were used for segmentation of the hippocampus. The brain extractions were manually checked for quality control. This revealed that five participants had unsatisfactory extractions, which was addressed by applying custom extraction parameters and rerunning BET until we yielded satisfactory skull-stripped brain extractions. Manual extractions were checked and approved by two raters (TV and CP).

Volume

Segmentation of the bilateral hippocampus was performed using FIRST, a toolbox of FSL. FIRST performs registration, segmentation based on Bayesian appearance, and boundary corrections to produce segmented subcortical structures (Patenaude et al. 2011). Hippocampal extractions were verified visually and were not deemed satisfactory for one participant. While all other segmentations were performed on bias-corrected full T1 images, for the unsatisfactory segmentation, the pipeline was run again on the brain-extracted image, which produced a satisfactory subcortical segmentation of the structure. Hippocampal raw volumes were calculated using the fslstats tool. Hippocampal volume was normalised by dividing it by total intracranial volume as estimated from the skull stripped image.

Shape

As part of the FIRST pipeline, vertex analysis was also performed on the bilateral hippocampus to establish if BCS is a predictor for changes in the hippocampal shape. The standard procedure was implemented in FIRST, by which each structure was linearly registered (using 6 degrees of freedom) to the sample-specific average surface and mapped

in MNI space. Analysis was carried out using the Randomise pipeline in FSL, in which permutation-based non-parametric analysis with 10,000 permutations for each factor of interest testing was ran and corrected for multiple comparisons using threshold-free cluster enhancement (Smith and Nichols 2009). The correlational design matrix contained the factor of interest, BCS, and covariates of age and education. This procedure resulted in spatial maps showing local contractions and expansions of the structure (i.e., perpendicular displacement from the study-specific template average surface) of interest as a function of bilingualism.

The participant with unsatisfactory hippocampal segmentations from the complete T1 scan had to be excluded from shape analysis as it could not be included in the generation of the study-specific template of the hippocampus. Therefore, the study-specific template of the hippocampal vertices for the shape analysis was created without this participant. Note that the manual extraction of the hippocampal volume from the brain-extracted image was successful for this participant.

Volumetric analysis

We investigated whether degree of bilingualism predicts hippocampal volume beyond other demographic factors and memory performance. To do so, we used fixed effects of demographic measures, individual test scores from the NIH toolbox and ACE-III and the LSBQ BCS, and random effects of participant to build and compare several hierarchical linear mixed-effects model in an increasing order of complexity.¹

The FIRST tool, used to segment the structures of interest, provides two volumetric values—one for the left hippocampus, and one for the right. Mixed-effects models were run to evaluate hippocampal volume on either hemisphere. The initial model (Model 1) explains hippocampus volume as a function of age, education, hemisphere, and random effects of participant. The second model (Model 2) adds memory performance measures as independent variables to the model. The decision to include memory performance measures as predictors in this model (whereas, more typically, one would see brain measures as predictors for behaviour) was done to account for the individual variance in the hippocampal volume, shown to account for behavioural performance in other studies. In other words, like the demographic variables, memory performance is effectively acting

as a predictor of no interest. The third model (Model 3) introduces our main predictor of interest, the LSBQ BCS.² Akaike Information Criterion (AIC) was established for all models using the `anova()` function to determine the goodness of fit and choose the most appropriate model for the data.

Behavioural analysis

Pertaining to the second hypothesis, we aimed to explore if bilingualism as a continuous variable predicts memory performance when other variables, including age, education, and total hippocampal volume, are accounted for. This included running separate sets of hierarchical models for all three memory performance measures—NIH toolbox episodic memory score, NIH toolbox working memory score, and ACE-III memory score. The models (Models 4–7) were built in a similar manner to the volumetric analysis models of the hippocampus (Models 1–3), with the difference of exclusion of a fixed effect of hemisphere and random intercepts of participant. For this analysis, hippocampal volumes were summed across hemispheres and the total hippocampal volume was used as a predictor for memory performance. For model numbering purposes, models testing episodic memory were numbered by adding the letter E, working memory models—letter W, and ACE-III memory composite—letter C.

In variants of Model 4, each memory measure as a dependent variable was predicted by age and education as independent variables. In the following step, variants of Model 5, total hippocampal volume was added to the list of independent variables. LSBQ BCS was added as an independent predictor in variants of Model 6. Finally, to test if bilingualism interacts with the hippocampal volume to impact memory in variants of Model 7, we explored the interaction of the hippocampal volume and BCS. This model was built by expanding model 6 with an age by bilingualism composite score interaction term.

All variants of models 4–7 were checked for assumptions for linear regressions using the `gvlma` package (Peña and Slate 2006). They were met for models testing episodic memory performance as the dependent variable. However, one or more assumptions were violated for models testing working memory and ACE-III memory domain score,

¹ Given the participant age, we did not use age of acquisition as a reliable predictor as many bilingual individuals reported learning English at school, which may not accurately portray the actual engagement in English use across the lifespan and at the time of testing.

² A version of Model 3 was also envisaged with L2 home and L2 social use scores as measures of bilingualism, and the interaction of two, instead of the LSBQ composite score. However, as these scores heavily contribute to the LSBQ composite score (L2 Home and L2 Social scores have 33% and 30% weighting in the BCS calculation), and they were highly correlated between them ($\text{cor}=0.95$; $p<0.001$), it was deemed inappropriate to include these scores as separate predictors in the same model as they introduce multicollinearity issues.

suggesting that results of these regression models may not be accurate.

Results

Neuroimaging results

Volume

Hierarchical mixed-effects models were used to investigate whether the observed increased hippocampal volumes can be predicted by the amount of bilingual experience (see Table 3). Results from Model 1 revealed a trending negative effect of age, such that with increasing age, the observed hippocampus volumes became smaller, and a trending positive effect of education where higher educational attainment predicts higher hippocampal volume. No significant effects of hemisphere were observed.

Results from Model 2 revealed that age and education became significant contributors to the hippocampal volume, whereas hemisphere was not. Moreover, performance in the episodic memory task of the NIH toolbox correlated negatively with hippocampal volume, whereas performance in the working memory task of the NIH toolbox and overall composite memory performance were not significantly associated with hippocampal volume.

Finally, Model 3 revealed that, while the effects of education, and episodic memory performance remained significant, BCS also emerged as a unique contributor to the hippocampal volume, with higher BCS being positively associated with hippocampal volume (see Table 3). Adding BCS to the model increased the marginal R^2 value from 0.225 to 0.28 and AIC decreased from -1299.1 to -1302.2 when comparing model 3 to model 2, indicating better explanatory power of the model by including BCS as a predictor. The lowest AIC indicating the best model fit for the data was for the most complex model (Model 3).

A version of model 3 (Model 3b) was also run on the bilingual subsample, to test the claim that higher amount of bilingual experiences correlates with change in neural anatomy. Within the bilingual participant group, the effect of LSBQ BCS was trending at $p = 0.0514$. This is not necessarily an unexpected outcome, as the power was drastically reduced, by including only 23 participants in this analysis, and LSBQ captures scores across the spectrum of *-lingualism*. No other predictor but episodic memory performance ($p < 0.01$) was significant in this version of the model (see Table 4).

To test whether bilingualism affects the hippocampus specifically, we ran a control model (Model 3c) with all predictors remaining as independent variables but substituting hippocampal volume with normalised brainstem

Table 3 Hippocampal volume model comparison (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$)

Predictors	Model 1			Model 2			Model 3		
	Estimates	SE	Std. Beta	Statistic	p	Estimates	SE	Std. Beta	Statistic
Intercept	0.003***	0.000	-0.071	6.143	< 0.001	0.003***	0.001	-0.071	4.429
Age	-0.000	0.000	-0.232	-1.779	0.082	-0.000*	0.000	-0.239	-1.782
Education	0.000	0.000	0.229	1.752	0.087	0.000*	0.000	0.275	2.213
Hemisphere	0.000	0.000	0.143	1.293	0.202	0.000	0.000	0.143	1.293
ACE-III memory						0.000	0.000	0.136	1.090
NIH-TB episodic memory						-0.000*	0.000	-0.302	-2.233
NIH-TB working memory						-0.000	0.000	-0.127	-1.056
LSBQ BCS						0.000*	0.000	0.282	2.131
Observations	96					96			
Marginal R^2 /Conditional R^2	0.133/0.719					0.225/0.729			
AIC	-1297.5					-1299.1			

The significance level for bold should be defined at the level of $p < 0.05$

Table 4 Model 3b (Model 3 on the bilingual subsample) (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$)

Predictors	Model 3b				
	Estimates	Std. Error	Std. Beta	Statistic	p
Intercept	0.003*	0.001	− 0.153	2.593	0.020
Age	− 0.000	0.000	− 0.255	− 1.120	0.279
Education	0.000	0.000	0.238	1.276	0.220
Hemisphere	0.000	0.000	0.306	1.617	0.120
ACE-III memory	0.000	0.000	0.099	0.554	0.588
NIH-TB episodic memory	− 0.000*	0.000	− 0.591	− 2.715	0.015
NIH-TB working memory	0.000	0.000	0.043	0.239	0.814
Bilingualism Composite Score	0.000	0.000	0.370	2.105	0.051
Observations	46				
Marginal R^2 /conditional R^2	0.303/0.651				

The significance level for bold should be defined at the level of $p < 0.05$

volume as the dependent variable. Brainstem was chosen as a comparison variable to the hippocampus as there is no theoretical reason to believe that bilingualism has any effect on brainstem volume. The results showed that, unlike for the hippocampal volume, bilingualism did not significantly predict brainstem volume, indicative of specificity for bilingualism effects to hippocampal volume. Statistics for this model are reported in the Supplemental material (Table S1).

Shape

Vertex analysis revealed no significant local expansions or contractions of the bilateral hippocampus as a function of BCS, thresholded at $p = < 0.05$.

Behavioural results

As no effect of hemisphere emerged in the first set of models, we summed the left and right hippocampal volumes and used total hippocampal volume as a predictor in the sets of models testing behavioural performance across memory domains. For the NIH toolbox working memory task, hierarchical regressions showed that none of the predictors (age, education, total hippocampal volume, BCS, or BCS by hippocampal volume interaction) significantly explained working memory performance (see Models 4W, 5W, 6W, 7W, Table 5).

For the NIH toolbox episodic memory task, age was a significant predictor in the Model 4E and remained a significant predictor in Models 5E, 6E, and 7E, so that with increased age, episodic memory performance is subject to decline. Education was a significant positive predictor in Models 5E, 6E, and 7E. Hippocampal volume also predicted episodic memory performance in Model 6E and 7E; however, the

relationship was negative. In Model 6E, BCS did not significantly contribute to episodic memory performance. From all four episodic memory models, Model 6E was also the best fit for data with an adjusted R^2 of 0.239 and the lowest AIC indicating the best fit (see Table 6).

For the ACE-III cognition battery memory domain, across all four models, no independent variables significantly predicted composite memory performance, apart from a significant main effect of sex in Models 4C and 5C, a trend for BCS in Model 6C ($p = 0.078$), suggesting that higher BCS might predict better performance in the ACE-III memory domain (see Table 7).

Unlike the models explaining the volumetric variation of the hippocampus as a result of demographic variables, memory performance, and bilingualism, the linear regression models explaining the variance in memory performance were not a good fit for the data. In all cases, model comparison revealed the increasingly complex models not to improve their explanatory power over the data. The only exception to this were models explaining NIH episodic memory scores as a function of the above-described IVs, where most complex model offered a marginal improvement over the simpler models ($p = 0.064$). Therefore, only the episodic memory performance can be measured as a function of age, education, hippocampal volume, and bilingualism. See hierarchical regression model comparison for all three memory scores in Tables 5, 6, 7.

No significant interaction of bilingualism and hippocampal volume was revealed on either memory measurement.

Discussion

In the present study, we examined the effects bilingualism might have on the ageing brain with a particular focus on the hippocampus and related cognitive abilities. The

NIH toolbox working memory

Performance in the NIH toolbox working memory task (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$)
The significance level for bold should be defined at the level of $p < 0.05$

The particular age range of our participants and the cross-sectional design of the study do not allow to differentiate between the above scenarios. Nevertheless, evidence from this exact pivotal point in cognitive ageing might prove useful in explicating effects of bilingualism later in life and/or under pathological neurodegeneration (Berkes et al. 2020;

Table 6 Behavioural hierarchical regression

NIH toolbox episodic memory																
Predictors	Model 4E				Model 5E				Model 6E				Model 7E			
	Estimates	Std. Error	Std. Beta	Statistic	<i>p</i>	Estimates	Std. Error	Std. Beta	Statistic	<i>p</i>	Estimates	Std. Error	Std. Beta	Statistic	<i>p</i>	
Intercept	109.970***	16.733	0.012	6.572	<0.001	140.677***	22.427	0.061	6.273	<0.001	142.013***	22.150	0.052	6.411	<0.001	
Age	−0.435*	0.178	−0.338	−2.435	0.019	−0.527**	0.179	−0.410	−2.944	0.005	−0.453*	0.184	−0.353	−2.467	0.018	
Sex	−0.402	3.410	−0.033	−0.118	0.907	−2.016	3.401	−0.163	−0.593	0.556	−1.705	3.363	−0.138	−0.507	0.615	
Education	0.832	0.503	0.231	1.655	0.105	1.111*	0.507	0.308	2.191	0.034	1.077*	0.501	0.298	2.150	0.037	
Hippocampal volume						−5831.010	2944.427	−0.289	−1.980	0.054	−7100.063*	3031.373	−0.352	−2.342	0.024	
LSBQ BCS											0.224	0.152	0.215	1.469	0.149	
Hippocampal volume × LSBQ BCS																
Observations	48					48					48					
<i>R</i> ² / <i>R</i> ² adjusted	0.206/0.152					0.273/0.205					0.308/0.226					
AIC	375.628					373.439					373.035					

Performance in the NIH episodic memory task (**p* < 0.05; ***p* < 0.01; ****p* < 0.001)

The significance level for bold should be defined at the level of *p* < 0.05

Costumero et al. 2020; Duncan et al. 2018). Most notably, our finding that individual-level engagement with bilingual experiences can affect the hippocampus structurally follows from similar findings in younger bilinguals (DeLuca et al. 2019b). Importantly, the present study constitutes the first piece of evidence that brain reserves specifically in older bilinguals are modulated by individual-level factors related to how one interacts with their languages. This alone is an important finding, because it clarifies the confines and parameters under which effects of bilingualism are likely to take place, reasonable to predict and worthy of serious consideration to be promoted as best practice for amelioration of age-related decline and neurodegeneration (e.g., Voits et al. 2020).

Hypothesis two related to effects bilingualism might have on the performance in cognitive domains typically associated with the hippocampus—most notably, episodic memory. It was hypothesised that if bilingualism had a measurable effect on hippocampal volume, positive behavioural effects would likely co-occur. Episodic and working memory performance in our samples was tested with three separate tasks. With the potential effects of demographic factors, such as age and educational attainment, as well as hippocampal volume all accounted for, bilingualism did not emerge as a significant predictor for memory performance across any of the tests we administered. Thus, hypothesis two was disconfirmed. In fact, an unexpected finding resulted: our data show a negative relationship between hippocampal volume and episodic memory performance. This is especially surprising given that positive associations of episodic memory and the hippocampus have been widely reported in adjacent literatures (Anand and Dhikav 2012; O'Shea et al. 2016). A potential explanation might stem from the fact that episodic memory performance is not uniquely reliant on hippocampal volume. Rather it is a network of cortical, subcortical, and medial temporal lobe structures that work in tandem (Dickerson and Eichenbaum 2010). Although the hippocampus plays a prominent role in this network, one might need to investigate the structural integrity of this network as a whole, which is beyond the scope of the present paper. In any case, a positive brain-to-behaviour relationship is intuitive, often empirically shown and theoretically reasonable. Thus, the general asymmetry we report is at first glance perplexing. We now turn to ponder how to best make sense of these juxtaposed outcomes.

Why the discrepancy between the effects of bilingualism on brain structure and cognitive performance? One explanation could be the average age of our participants being only 62 years. This puts them on the 'younger' end of the ageing spectrum. While some ageing processes may have already begun, these participants are still cognitively healthy with no signs of memory impairment, attested by the near-ceiling score in the ACE-III memory domain. Nevertheless, our

Table 7 Behavioural hierarchical regression

ACE-III composite memory score																					
Predictors	Model 4C			Model 5C			Model 6C			Model 7C			Statistic	p							
	Estimates	Std. Error	Std. Beta	Statistic	p	Estimates	Std. Error	Std. Beta	Statistic	p	Estimates	Std. Error			Std. Beta						
Intercept	24.612***	3.070	0.253	8.016	<0.001	22.730***	4.279	0.236	5.312	<0.001	23.044***	4.169	0.223	5.528	<0.001	23.024***	4.217	0.240	5.460	<0.001	
Age	-0.009	0.033	-0.040	-0.269	0.789	-0.003	0.034	-0.014	-0.092	0.927	0.014	0.035	0.064	0.411	0.683	0.015	0.035	0.068	0.431	0.668	
Sex	-1.449 *	0.626	-0.675	-2.315	0.025	-1.350 *	0.649	-0.629	-2.080	0.044	-1.277	0.633	-0.595	-2.017	0.050	-1.275	0.640	-0.594	-1.992	0.053	
Education	0.041	0.092	0.066	0.445	0.659	0.024	0.097	0.038	0.248	0.806	0.016	0.094	0.025	0.169	0.867	0.013	0.096	0.021	0.134	0.894	
Hippocampal volume						357.382	561.741	0.102	0.636	0.528	58.882	570.554	0.017	0.103	0.918	67.749	578.040	0.000	0.117	0.907	
LSBQ BCS											0.053	0.029	0.291	1.835	0.074	0.123	0.277	0.307	0.445	0.659	
Hippocampal volume x LSBQ BCS																	-13.275	51.732	-0.045	-0.257	0.799
Observations	48					48					48										
R ² /R ² adjusted	0.112/0.052					0.120/0.039					0.186/0.089										
AIC	212.852					214.4021					212.6991										

Performance in the ACE-III memory component (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$)

The significance level for bold should be defined at the level of $p < 0.05$

structural findings indicate that the processes that underlie the building of a brain reserve are already in action, but without measurable equivalents in behaviour. This pattern is reminiscent of evidence, suggesting that the mapping of behaviour to brain *function* is not always straightforward, at least in healthy populations (Abutalebi et al. 2012; DeLuca et al. 2020).

If on the right track, our participants, whether bilingual or monolingual, are not deep in the process of cognitive ageing. As a result, the behavioural asymmetry evidence could help to specify which of the scenarios articulated above for the monolingual/bilingual difference in hippocampal volume is more likely to be on target. It would follow from this line of reasoning that there is a volumetric increase for bilinguals (on a sliding scale relative to linguistic experience), prior to any significant cognitive ageing effects. If so, this would ostensibly leave the brain-behaviour performance asymmetry perplexing only to the extent that a mapping between hippocampal volume and increased memory performance must follow. On second thought, however, we submit that the asymmetry is not overly surprising nor concerning. In our data, generalised memory task performance is at or near ceiling across the board, which alone may indicate the lack of significant cognitive ageing in our participants. Given the very high performance, it is reasonable to assume that the granularity of the memory tasks is simply not sufficient to capture potentially underlying differences in memory represented by increased volume in the bilinguals. When study participants perform at ceiling on a given behavioural task, it is functionally impossible to further test whether all individuals are equipotential for the construct of interest. In other words, we cannot preclude that bilinguals with increased exposure/engagement relative to others (monolinguals or bilinguals with less engagement) do not have better memory resources overall. Rather what we know for sure is that everyone has sufficient memory resources to perform these specific tasks at ceiling. The behavioural tasks used in this study can be viewed, then, as a limitation. Behavioural effects of bilingualism have been shown to manifest (or not) depending on task difficulty (Costa et al. 2009). And so, future studies should employ more difficult cognitive tasks where individuals would be less likely to score at ceiling.

An inspirational study for the present one was that of Schroeder and Marian (2012) in which a positive correlation between bilingualism and memory performance was shown. However, the bilinguals in that study had a mean age of 80+. As age increases, the effects of cognitive ageing and neuropathology can escalate exponentially (Fox and Schott 2004). Taken together, it could be the case that our participants, while older, are not old enough as in Schroeder and Marian (2012) to exhibit behavioural differences in episodic memory, at least on the specific tasks we used. All things being equal, we might expect that the bilingual participants in Schroeder and Marian

(2012) to have similar or even more signs of neural atrophy to the hippocampus despite greater behavioural performance compared to their controls. Why? Because in their age range, one would expect that accrued neural reserve is being or has been exhausted and the compensation processes for neural atrophy (behaviour task corollaries of cognitive reserve) play a principal role. This, of course, is an empirical question—and a future avenue of research—as the relevant data do not exist. To the extent that using (up) of neural reserve precedes, even if overlapping to some degree, the behavioural effects of cognitive reserve, the present data would add support to this argumentation. A clear example of this is a recent study where monolinguals and bilinguals were matched on brain health (unlike the more commonly used matching on cognitive performance), which showed a bilingualism-related maintenance of cognitive status at equal levels of brain decline in ageing individuals (Berkes et al. 2021). In sum, we interpret the observed volumetric brain evidence an index of brain reserve.

Finally, the volumetric changes across the participant sample did not translate into significant effects on hippocampal shape. This lacking relationship is more challenging to interpret, not least as it leaves some questions for understanding the relationship between volume and shape as they are assessed by our tools. Regardless, this does not seem to be unprecedented within the relevant literature where the few available studies have, similarly to our data, reported effects on one metric only (volume or shape), but not both (DeLuca et al. 2019b; Li et al. 2017; Mårtensson et al. 2012).

Our findings call for further and more focused investigations on bilingual engagement effects on the ageing brain, and in particular on age ranges similar to ours, where the first signs of cognitive decline might emerge. Of course, the ultimate goal in this programme is to reveal the exact mechanisms of how bilingual experience and increased executive control demands impact episodic memory; however, addressing this properly requires much more research and sits beyond the scope of this study. Longitudinal designs would enable further, and more precise examination of onset and trajectory of any relationship bilingualism has in exponents of cognitive ageing as well as their underlying mechanisms. Crucially, focused studies similar to this one are required with pathologically ageing populations too, to add to a small but growing literature that will help us better understand the potential clinical implications both in healthy and pathological ageing (Voits et al. 2020). Moreover, our results suggest that given the observed decoupling of brain structure and behaviour, augmenting studies of behavioural task performance with methodologies that directly look under the proverbial hood simultaneously are especially welcome to assess the effects of bilingualism on the brain.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s00429-021-02436-z>.

Funding Open Access funding provided by UiT The Arctic University of Norway. The overall project received funding from the Centre for Integrative Neuroscience and Neurodynamics as well as the School of Psychology and Clinical Language Sciences at the University of Reading. Toms Voits is funded by the AcqVA Aurora Centre grant and Jason Rothman is funded by the same grant as well as a Tromsø Forskningsstiftelse (TFS) grant.

Availability of data Data will be made available on Openneuro and/or OSF and accession numbers provided on request.

Code availability Code will be made available on OSF.

Declarations

Conflict of interest None.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abutalebi J, Della Rosa PA, Green DW, Hernandez M, Scifo P, Keim R, Cappa SF, Costa A (2012) Bilingualism tunes the anterior cingulate cortex for conflict monitoring. *Cereb Cortex* 22(9):2076–2086. <https://doi.org/10.1093/cercor/bhr287>
- Alladi S, Bak TH, Duggirala V, Surampudi B, Shailaja M, Shukla AK, Chaudhuri JR, Kaul S (2013) Bilingualism delays age at onset of dementia, independent of education and immigration status. *Neurology* 81:1938–1944. <https://doi.org/10.1212/01.WNL.0000437347.92583.fc>
- Anand K, Dhikav V (2012) Hippocampus in health and disease: an overview. *Ann Indian Acad Neurol* 15(4):239–246. <https://doi.org/10.4103/0972-2327.104323>
- Anderson JAE, Grundy JG, De Frutos J, Barker RM, Grady C, Bialystok E (2018a) Effects of bilingualism on white matter integrity in older adults. *NeuroImage* 167:143–150. <https://doi.org/10.1016/j.neuroimage.2017.11.038>
- Anderson JAE, Mak L, Keyvani Chahi A, Bialystok E (2018b) The language and social background questionnaire: Assessing degree of bilingualism in a diverse population. *Behav Res Methods* 50(1):250–263. <https://doi.org/10.3758/s13428-017-0867-9>
- Anderson JAE, Hawrylewicz K, Grundy JG (2020) Does bilingualism protect against dementia? A meta-analysis. *Psychon Bull Rev*. <https://doi.org/10.3758/s13423-020-01736-5>
- Apostolova LG, Dutton RA, Dinov ID, Hayashi KM, Toga AW, Cummings JL, Thompson PM (2006) Conversion of mild cognitive impairment to alzheimer disease predicted by hippocampal atrophy maps. *Arch Neurol* 63(5):693–699. <https://doi.org/10.1001/archneur.63.5.693>

- Bak TH (2016) Cooking pasta in La Paz. *Linguist Approaches Biling* 6(5):699–717. <https://doi.org/10.1075/lab.16002.bak>
- Bellander M, Berggren R, Mårtensson J, Brehmer Y, Wenger E, Li TQ, Bodammer NC, Shing YL, Werkle-Bergner M, Lövdén M (2016) Behavioral correlates of changes in hippocampal gray matter structure during acquisition of foreign vocabulary. *Neuroimage* 131:205–213. <https://doi.org/10.1016/j.neuroimage.2015.10.020>
- Bellmund JLS, Gärdenfors P, Moser EI, Doeller CF (2018) Navigating cognition: spatial codes for human thinking. *Science* (New York, N Y). <https://doi.org/10.1126/science.aat6766>
- Berkes M, Bialystok E, Craik FIM, Troyer A, Freedman M (2020) Conversion of mild cognitive impairment to alzheimer disease in monolingual and bilingual patients. *Alzheimer Dis Assoc Disord* 00(00):1–6. <https://doi.org/10.1097/WAD.0000000000000373>
- Berkes M, Calvo N, Anderson JAE, Bialystok E (2021) Poorer clinical outcomes for older adult monolinguals when matched to bilinguals on brain health. *Brain Struct Funct*. <https://doi.org/10.1007/s00429-020-02185-5>
- Bettio LEB, Rajendran L, Gil-Mohapel J (2017) The effects of aging in the hippocampus and cognitive decline. *Neurosci Biobehav Rev* 79:66–86. <https://doi.org/10.1016/j.neubiorev.2017.04.030>
- Bialystok E (2017) The bilingual adaptation: How minds accommodate experience. *Psychol Bull* 143(3):233–262. <https://doi.org/10.1037/bul0000099>
- Bialystok E, Craik FIM, Klein R, Viswanathan M (2004) Bilingualism, aging, and cognitive control: evidence from the simon task. *Psychol Aging* 19(2):290–303. <https://doi.org/10.1037/0882-7974.19.2.290>
- Bialystok E, Craik FIM, Freedman M (2007) Bilingualism as a protection against the onset of symptoms of dementia. *Neuropsychologia* 45(2):459–464. <https://doi.org/10.1016/j.neuropsychologia.2006.10.009>
- Bird CM, Burgess N (2008) The hippocampus and memory: insights from spatial processing. *Nat Rev Neurosci* 9(3):182–194. <https://doi.org/10.1038/nrn2335>
- Borsa VM, Perani D, Della Rosa PA, Videsott G, Guidi L, Weekes BS, Franceschini R, Abutalebi J (2018) Bilingualism and healthy aging: aging effects and neural maintenance. *Neuropsychologia* 111(January):51–61. <https://doi.org/10.1016/j.neuropsychologia.2018.01.012>
- Breitenstein C, Jansen A, Deppe M, Foerster AF, Sommer J, Wolbers T, Knecht S (2005) Hippocampus activity differentiates good from poor learners of a novel lexicon. *Neuroimage* 25(3):958–968. <https://doi.org/10.1016/j.neuroimage.2004.12.019>
- Bruno D, Vignaga SS (2019) Addenbrooke's cognitive examination III in the diagnosis of dementia: a critical review. *Neuropsychiatr Dis Treat* 15:441–447. <https://doi.org/10.2147/NDT.S151253>
- Cabeza R, Albert M, Belleville S, Craik FIMM, Duarte A, Grady CL, Lindenberger U, Nyberg L, Park DC, Reuter-Lorenz PA, Rugg MD, Steffener J, Rajah MN (2018) Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. *Nat Rev Neurosci* 19(11):701–710. <https://doi.org/10.1038/s41583-018-0068-2>
- Calabria M, Hernández M, Cattaneo G, Suades A, Serra M, Juncadella M, Reñé R, Sala I, Lleó A, Ortiz-Gil J, Ugas L, Ávila A, Ruiz IG, Ávila C, Costa A (2020) Active bilingualism delays the onset of mild cognitive impairment. *Neuropsychologia*. <https://doi.org/10.1016/j.neuropsychologia.2020.107528>
- Costa A, Hernández M, Costa-Faidella J, Sebastián-Gallés N (2009) On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition* 113(2):135–149. <https://doi.org/10.1016/j.cognition.2009.08.001>
- Costumero V, Marin-Marín L, Calabria M, Belloch V, Escudero J, Baquero M, Hernandez M, Ruiz de Miras J, Costa A, Parcet M-A, Ávila C (2020) A cross-sectional and longitudinal study on the protective effect of bilingualism against dementia using brain atrophy and cognitive measures. *Alzheimer's Res Ther* 12(1):11. <https://doi.org/10.1186/s13195-020-0581-1>
- Darwish H, Farran N, Assaad S, Chaaya M (2018) Cognitive reserve factors in a developing country: education and occupational attainment lower the risk of dementia in a sample of lebanese older adults. *Front Aging Neurosci* 10:1–10. <https://doi.org/10.3389/fnagi.2018.00277>
- de Bruin A (2019) Not all bilinguals are the same: a call for more detailed assessments and descriptions of bilingual experiences. *Behav Sci* 9(3):1–13. <https://doi.org/10.3390/BS9030033>
- De Cat C, Gusnanto A, Serratrice L (2018) Identifying a threshold for the executive function advantage in bilingual children. *Stud Second Lang Acquis* 40(1):119–151. <https://doi.org/10.1017/S0272263116000486>
- DeLuca V, Rothman J, Bialystok E, Pliatsikas C (2019a) Redefining bilingualism as a spectrum of experiences that differentially affects brain structure and function. *Proc Natl Acad Sci* 116(15):7565–7574. <https://doi.org/10.1073/pnas.1811513116>
- DeLuca V, Rothman J, Pliatsikas C (2019) Linguistic immersion and structural effects on the bilingual brain: a longitudinal study. *Biling Lang Cogn* 22(5):1160–1175. <https://doi.org/10.1017/S1366728918000883>
- DeLuca V, Rothman J, Bialystok E, Pliatsikas C (2020) Duration and extent of bilingual experience modulate neurocognitive outcomes. *NeuroImage* 204:116222. <https://doi.org/10.1016/j.neuroimage.2019.116222>
- Dickerson BC, Eichenbaum H (2010) The episodic memory system: neurocircuitry and disorders. *Neuropsychopharmacology* 35(1):86–104. <https://doi.org/10.1038/npp.2009.126>
- Duncan HD, Nikelski J, Pilon R, Steffener J, Chertkow H, Phillips NA (2018) Structural brain differences between monolingual and multilingual patients with mild cognitive impairment and Alzheimer disease: evidence for cognitive reserve. *Neuropsychologia* 109:270–282. <https://doi.org/10.1016/j.neuropsychologia.2017.12.036>
- Erickson KI, Voss MW, Prakash RS, Basak C, Szabo A, Chaddock L, Kim JS, Heo S, Alves H, White SM, Wojcicki TR, Mailley E, Vieira VJ, Martin SA, Pence BD, Woods JA, McAuley E, Kramer AF (2011) Exercise training increases size of hippocampus and improves memory. *Proc Natl Acad Sci USA* 108(7):3017–3022. <https://doi.org/10.1073/pnas.1015950108>
- Erten-Lyons D, Woltjer RL, Dodge H, Nixon R, Vorobik R, Calvert JF, Leahy M, Montine T, Kaye J (2009) Factors associated with resistance to dementia despite high Alzheimer disease pathology. *Neurology* 72(4):354–360. <https://doi.org/10.1212/01.wnl.0000341273.18141.64>
- Farokhian F, Yang C, Beheshti I, Matsuda H, Wu S (2017) Age-related gray and white matter changes in normal adult brains. *Aging and Disease* 8(6):899–909. <https://doi.org/10.14336/AD.2017.0502>
- Firth J, Stubbs B, Vancampfort D, Schuch F, Lagopoulos J, Rosenbaum S, Ward PB (2018) Effect of aerobic exercise on hippocampal volume in humans: a systematic review and meta-analysis. *NeuroImage* 166:230–238. <https://doi.org/10.1016/j.neuroimage.2017.11.007>
- Fjell AM, Walhovd KB, Fennema-Notestine C, McEvoy LK, Hagler DJ, Holland D, Brewer JB, Dale AM (2009) One-year brain atrophy evident in healthy aging. *J Neurosci* 29(48):15223–15231. <https://doi.org/10.1523/JNEUROSCI.3252-09.2009>
- Fotuhi M, Do D, Jack C (2012) Modifiable factors that alter the size of the hippocampus with ageing. *Nat Rev Neurol* 8(4):189–202. <https://doi.org/10.1038/nrneurol.2012.27>
- Foubert-Samier A, Catheline G, Amieva H, Dilharreguy B, Helmer C, Allard M, Dartigues JF (2012) Education, occupation, leisure activities, and brain reserve: a population-based study. *Neurobiol*

- Aging 33(2):423.e15–423.e25. <https://doi.org/10.1016/j.neurobiolaging.2010.09.023>
- Fox NC, Schott JM (2004) Imaging cerebral atrophy: normal ageing to Alzheimer's disease. *The Lancet* 363(9406):392–394. [https://doi.org/10.1016/S0140-6736\(04\)15441-X](https://doi.org/10.1016/S0140-6736(04)15441-X)
- Fraser MA, Shaw ME, Cherbuin N (2015) A systematic review and meta-analysis of longitudinal hippocampal atrophy in healthy human ageing. *Neuroimage* 112:364–374. <https://doi.org/10.1016/j.neuroimage.2015.03.035>
- Giorgio A, Santelli L, Tomassini V, Bosnell R, Smith S, De Stefano N, Johansen-Berg H (2010) Age-related changes in grey and white matter structure throughout adulthood. *Neuroimage* 51(3):943–951. <https://doi.org/10.1016/j.neuroimage.2010.03.004>
- Gold BT, Kim C, Johnson NF, Kryscio RJ, Smith CD (2013) Lifelong bilingualism maintains neural efficiency for cognitive control in aging. *J Neurosci* 33(2):387–396. <https://doi.org/10.1523/jneurosci.3837-12.2013>
- Gorbach T, Pudas S, Lundquist A, Orädd G, Josefsson M, Salami A, de Luna X, Nyberg L (2017) Longitudinal association between hippocampus atrophy and episodic-memory decline. *Neurobiol Aging* 51:167–176. <https://doi.org/10.1016/j.neurobiolaging.2016.12.002>
- Grundy JG (2020) The effects of bilingualism on executive functions: an updated quantitative analysis. *J Cult Cogn Sci*. <https://doi.org/10.1007/s41809-020-00062-5>
- Grundy JG, Anderson JAE, Bialystok E (2017) Neural correlates of cognitive processing in monolinguals and bilinguals. *Ann N Y Acad Sci*. <https://doi.org/10.1111/nyas.13333>
- Gullifer JW, Chai XJ, Whitford V, Pivneva I, Baum S (2018) Neuropsychologia Bilingual experience and resting-state brain connectivity: impacts of L2 age of acquisition and social diversity of language use on control networks. *Neuropsychologia* 117(May):123–134. <https://doi.org/10.1016/j.neuropsychologia.2018.04.037>
- Hayakawa S, Marian V (2019) Consequences of multilingualism for neural architecture. *Behav Brain Funct* 15(1):1–24. <https://doi.org/10.1186/s12993-019-0157-z>
- Hötting K, Röder B (2013) Beneficial effects of physical exercise on neuroplasticity and cognition. *Neurosci Biobehav Rev* 37(9):2243–2257. <https://doi.org/10.1016/j.neubiorev.2013.04.005>
- Hsieh S, Schubert S, Hoon C, Mioshi E, Hodges JR (2013) Validation of the Addenbrooke's Cognitive Examination III in Frontotemporal Dementia and Alzheimer's Disease. *Dement Geriatr Cogn Disord* 36(3–4):242–250. <https://doi.org/10.1159/000351671>
- Lehtonen M, Soveri A, Laine A, Järvenpää J, de Bruin A, Antfolk J (2018) Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychol Bull* 144(4):394–425. <https://doi.org/10.1037/bul0000142>
- Leivada E, Westergaard M, Duñabeitia JA, Rothman J (2021) On the phantom-like appearance of bilingualism effects on neurocognition: (How) should we proceed? *Biling Lang Cogn* 24(1):197–210. <https://doi.org/10.1017/S1366728920000358>
- Li L, Abutalebi J, Emmorey K, Gong G, Yan X, Feng X, Zou L, Ding G (2017) How bilingualism protects the brain from aging: Insights from bimodal bilinguals. *Hum Brain Mapp* 38(8):4109–4124. <https://doi.org/10.1002/hbm.23652>
- Lövdén M, Schaefer S, Noack H, Bodammer NC, Kühn S, Heinze HJ, Düzel E, Bäckman L, Lindenberger U (2012) Spatial navigation training protects the hippocampus against age-related changes during early and late adulthood. *Neurobiol Aging* 33(3):620.e9–620.e22. <https://doi.org/10.1016/j.neurobiolaging.2011.02.013>
- Luk G, Bialystok E (2013) Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *J Cogn Psychol* 25(5):605–621. <https://doi.org/10.1080/20445911.2013.795574>
- Mann A, de Bruin A (2021) Bilingual language use is context dependent: using the Language and Social Background Questionnaire to assess language experiences and test-rest reliability. *Int J Biling Educ Biling*. <https://doi.org/10.1080/13670050.2021.1988049>
- Mårtensson J, Eriksson J, Christian N, Lindgren M, Johansson M, Nyberg L, Lövdén M, Bodammer NC, Lindgren M, Johansson M, Nyberg L, Lövdén M, Christian N, Lindgren M, Johansson M, Nyberg L, Lövdén M (2012) Growth of language-related brain areas after foreign language learning. *Neuroimage* 63(1):240–244. <https://doi.org/10.1016/j.neuroimage.2012.06.043>
- Mukadam N, Sommerlad A, Livingston G (2017) The relationship of bilingualism compared to monolingualism to the risk of cognitive decline or dementia: a systematic review and meta-analysis. *J Alzheimer's Dis* 58(1):45–54. <https://doi.org/10.3233/JAD-170131>
- Nee D, D'Esposito M (2015) Working Memory. *Brain Mapping*. Elsevier, pp 589–595
- Nyberg L, Salami A, Andersson M, Eriksson J, Kalpouzos G, Kauppi K, Lind J, Pudas S, Persson J, Nilsson LG (2010) Longitudinal evidence for diminished frontal cortex function in aging. *Proc Natl Acad Sci USA* 107(52):22682–22686. <https://doi.org/10.1073/pnas.1012651108>
- O'Shea A, Cohen RA, Porges EC, Nissim NR, Woods AJ (2016) Cognitive aging and the hippocampus in older adults. *Front Aging Neurosci* 8:1–8. <https://doi.org/10.3389/fnagi.2016.00298>
- Olsen RK, Pangelinan MM, Bogulski C, Chakravarty MM, Luk G, Grady CL, Bialystok E (2015) The effect of lifelong bilingualism on regional grey and white matter volume. *Brain Res* 1612:128–139. <https://doi.org/10.1016/j.brainres.2015.02.034>
- Peña EA, Slate EH (2006) Global validation of linear model assumptions. *J Am Stat Assoc* 101(473):341. <https://doi.org/10.1198/016214505000000637>
- Patenaude B, Smith SM, Kennedy DN, Jenkinson M (2011) A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* 56(3):907–922. <https://doi.org/10.1016/j.neuroimage.2011.02.046>
- Perneckzy R, Kempermann G, Korczyn AD, Matthews FE, Ikram MA, Scarmeas N, Chetelat G, Stern Y, Ewers M (2019) Translational research on reserve against neurodegenerative disease: Consensus report of the International Conference on Cognitive Reserve in the Dementias and the Alzheimer's Association Reserve, Resilience and Protective Factors Professional Interest Ar. *BMC Medicine* 17(1):1–15. <https://doi.org/10.1186/s12916-019-1283-z>
- Persson J, Nyberg L, Lind J, Larsson A, Nilsson LG, Ingvar M, Buckner RL (2006) Structure-function correlates of cognitive decline in aging. *Cereb Cortex* 16(7):907–915. <https://doi.org/10.1093/cercor/bhj036>
- Pliatsikas C (2020) Understanding structural plasticity in the bilingual brain: the dynamic restructuring model. *Biling Lang Cogn* 23(2):459–471. <https://doi.org/10.1017/S1366728919000130>
- Pliatsikas C, DeLuca V, Moschopoulou E, Saddy JD (2017) Immersive bilingualism reshapes the core of the brain. *Brain Struct Funct* 222(4):1785–1795. <https://doi.org/10.1007/s00429-016-1307-9>
- Pliatsikas C, DeLuca V, Voits T (2020) The many shades of bilingualism: language experiences modulate adaptations in brain structure. *Lang Learn* 70(S2):133–149. <https://doi.org/10.1111/lang.12386>
- Pliatsikas C, Pereira Soares SM, Voits T, Deluca V, Rothman J (2021) Bilingualism is a long-term cognitively challenging experience that modulates metabolite concentrations in the healthy brain. *Sci Rep* 11(1):1–12. <https://doi.org/10.1038/s41598-021-86443-4>
- Ritchie SJ, Hill WD, Marioni RE, Davies G, Hagenaars SP, Harris SE, Cox SR, Taylor AM, Corley J, Pattie A, Redmond P, Starr JM, Deary IJ (2019) Polygenic predictors of age-related decline in cognitive ability. *Mol Psychiatry*. <https://doi.org/10.1038/s41380-019-0372-x>

- Rönnlund M, Nyberg L, Bäckman L, Nilsson LG (2005) Stability, growth, and decline in adult life span development of declarative memory: cross-sectional and longitudinal data from a population-based study. *Psychol Aging* 20(1):3–18. <https://doi.org/10.1037/0882-7974.20.1.3>
- Schroeder SR, Marian V (2012) A bilingual advantage for episodic memory in older adults. *J Cogn Psychol* 24(5):591–601. <https://doi.org/10.1080/20445911.2012.669367>
- Smith SM, Nichols TE (2009) Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44(1):83–98. <https://doi.org/10.1016/j.neuroimage.2008.03.061>
- Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23(SUPPL. 1):208–219. <https://doi.org/10.1016/j.neuroimage.2004.07.051>
- Stern Y (2002) What is cognitive reserve? Theory and research application of the reserve concept. *J Int Neuropsychol Soc JINS* 8(3):448–460. <https://doi.org/10.1017/S1355617702813248>
- Stern Y, Arenaza-Urquijo EM, Bartrés-Faz D, Belleville S, Cantillon M, Chetelat G, Ewers M, Franzmeier N, Kempermann G, Kremen WS, Okonkwo O, Scarmeas N, Soldan A, Udeh-Momoh C, Valenzuela M, Vemuri P, Vuoksima E, Arenaza Urquijo EM, Bartrés-Faz D, Vuoksima E (2018) Whitepaper: defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimer's Dementia*. <https://doi.org/10.1016/j.jalz.2018.07.219>
- Sullivan MD, Prescott Y, Goldberg D, Bialystok E (2016) Executive control processes in verbal and nonverbal working memory: the role of aging and bilingualism. *Linguist Approaches Biling* 6(1–2):147–170. <https://doi.org/10.1075/lab.15056.sul>
- Sumowski JF, Rocca MA, Leavitt VM, Riccitelli G, Meani A, Comi G, Filippi M (2016) Reading, writing, and reserve: Literacy activities are linked to hippocampal volume and memory in multiple sclerosis. *Mult Scler* 22(12):1621–1625. <https://doi.org/10.1177/1352458516630822>
- Surraín S, Luk G (2019) Describing bilinguals: a systematic review of labels and descriptions used in the literature between 2005–2015. *Biling Lang Cogn* 22(2):401–415. <https://doi.org/10.1017/S1366728917000682>
- Ullman MT (2004) Contributions of memory circuits to language: the declarative/procedural model. *Cognition* 92(1–2):231–270. <https://doi.org/10.1016/j.cognition.2003.10.008>
- Valenzuela MJ, Sachdev P (2006) Brain reserve and dementia: a systematic review. *Psychol Med* 36(4):441–454. <https://doi.org/10.1017/S0033291705006264>
- Voits T, Pliatsikas C, Robson H, Rothman J (2020) Beyond Alzheimer's disease: can bilingualism be a more generalized protective factor in neurodegeneration? *Neuropsychologia* 147:107593. <https://doi.org/10.1016/j.neuropsychologia.2020.107593>
- Weintraub S, Dikmen SS, Heaton RK, Tulsky DS, Zelazo PD, Bauer PJ, Carlozzi NE, Slotkin J, Blitz D, Wallner-Allen K, Fox NA, Beaumont JL, Mungas D, Nowinski CJ, Richler J, Deocampo JA, Anderson JE, Manly JJ, Borosh B, Gershon RC (2013) Cognition assessment using the NIH Toolbox. *Neurology* 80:54–64. <https://doi.org/10.1212/WNL.0b013e3182872ded>
- Yaffe K, Fiocco AJ, Lindquist K, Vittinghoff E, Simonsick EM, Newman AB, Satterfield S, Rosano C, Rubin SM, Ayonayon HN, Harris TB (2009) Predictors of maintaining cognitive function in older adults: the Health ABC Study. *Neurology* 72(23):2029–2035. <https://doi.org/10.1212/WNL.0b013e3181a92c36>
- Yeung CM, St. John PD, Menec V, Tyas SL (2014) Is bilingualism associated with a lower risk of dementia in community-living older adults? cross-sectional and prospective analyses. *Alzheimer Dis Assoc Disord* 28(4):326–332. <https://doi.org/10.1097/WAD.000000000000019>
- Zahodne LB, Schofield PW, Farrell MT, Stern Y, Manly JJ (2014) Bilingualism does not alter cognitive decline or dementia risk among spanish-speaking immigrants. *Neuropsychology* 28(2):238–246. <https://doi.org/10.1037/neu0000014>

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