

Types of leisure activities differentially impact the function and structure of the brain

Michael R. Hess & Joukje M. Oosterman

Donders Institute for Brain, Cognition, and Behavior, Centre for Cognition, Radboud University, Nijmegen, The Netherlands

Leisure activity engagement has long been understood as an influential factor in the production and maintenance of cognitive performance. However, it remains unclear how to properly sub-group leisure activities so as to investigate and discern the differential cognitive impacts elicited by engaging in varying types of leisure activity. By using leisure activity data from the CRIq and grouping these data using factor-based analysis, I arrived at four unique leisure activity types, self-labeled: “responsibilities”, “entertainment”, “social/physical”, and “caretaking”. I then investigated the predictability of these groupings on behavioral and MRI measures by comparing group averages via statistical testing. Through these analyses, I found that this method of leisure activity grouping resulted in both significant cognitive performance and gray matter volume differences. This finding substantiates the need to further explore stricter categorical groupings of leisure activities in order to better understand their potential impact on the brain. In addition to this primary analysis, separating subject data by overall leisure activity engagement found multiple behavioral differences between higher and lower activity-engaging participants, but no structural brain differences were established in subsequent MRI analyses. An analysis of activities ranked by required cognitive effort found only one behavioral difference and analyses comparing activities categorized as “cognitive” or “social” resulted in no significant results.

Keywords: cognitive reserve, leisure activities, passive reserve, active reserve, cognitive trajectory

I. INTRODUCTION

Cognitive reserve (CR) is a theoretical construct describing the capacity and ability of the brain to function relatively well across an individual's adult lifespan despite ongoing age-related or disease-related atrophy. This idea of the brain possessing a resilience of functional ability derives from repeated observations of a mismatch between brain disease progression and the clinical manifestations of that pathology. In 1988, Katzman and colleagues surprisingly discovered cognitively well-functioning elderly women who, during examination of their brains upon death, were discovered to possess clinical features of Alzheimer's disease (AD). As well, their brains were found to weigh more due to containing a greater of neurons than those who had experienced AD symptomology. Katzman and colleagues hypothesized that these individuals may have avoided losses in functioning due to a larger reserve of excess neurons (Katzman, R. et al., 1988). From related lines of research, inter-individual variability in cognitive performance and brain pathology has increasingly become understood. While this variability can be attributed partly to genetics, a significant portion cannot. The environment the individual grows up in and the actions he partakes in are increasingly being understood to alter the cognitive trajectory across his lifespan by significantly influencing CR (Cheng, S.T., 2016; Mandolesi, L. et al., 2008).

To better understand CR and how it can be influenced by factors such as these, I will outline current models of reserve. CR can be broken down into two conceptual distinctions, a passive model of reserve and an active model of CR. In the passive model of reserve, known as brain reserve (BR), cognitive ability is maintained despite atrophy simply due to possessing enough additional neurons and their connections to compensate for the functional loss caused by the loss of those neurons. This redundancy of parallel neuronal connections provides back-up bridges between brain areas which can then be used if and when other connections are lost. Since this process requires no active engagement of the brain to occur, it is considered *passive* reserve.

The active model of reserve is conceptualized as the brain's ability to compensate for damage by enlisting the aid of pre-existing brain networks or by using alternative cognitive strategies. An illustrative example of this would be to consider the completion of a complex mathematical problem by a mathematician compared to someone less adept at math. The mathematician may have multiple strategies to completing the problem while the average person may have

only one. Thus, if one strategy is lost, the mathematician can still complete the problem while the other person has completely lost the ability. This redundancy of brain *strategies* is what defines the active model. For the purposes of brevity, these two conceptual models will both be considered components of overall CR.

These components are important measures in clinical cognitive science, and in fact have been implicated in predicting risk of incidental dementia and other age-related brain diseases, as well as delaying the time at which symptoms of brain diseases become noticeable. This occurs once the loss of neurons overtakes the individual's ability to functionally compensate for that loss. This point in time in which brain disease symptoms appear is commonly referred to as the "point of inflection", and is thought to be modulated by the individual's level of CR (Figure 1). Low reserve individuals exhibit symptomatology much earlier than high reserve individuals, however the point at which overall higher cognitive loss occurs appears temporally similar. This results in high reserve individuals displaying disease symptoms later, but experiencing a relatively faster rate of decrease in cognitive function once symptoms occur. One explanation for this phenomenon is that neuronal atrophy affects the brain network which mediates processes of CR functions at the same timepoint regardless of overall CR (Steffener, J., & Stern, Y., 2012). This means that despite the neuronal and network efficiency advantages present in high CR individuals, their ability to benefit from those advantages is eventually lost.

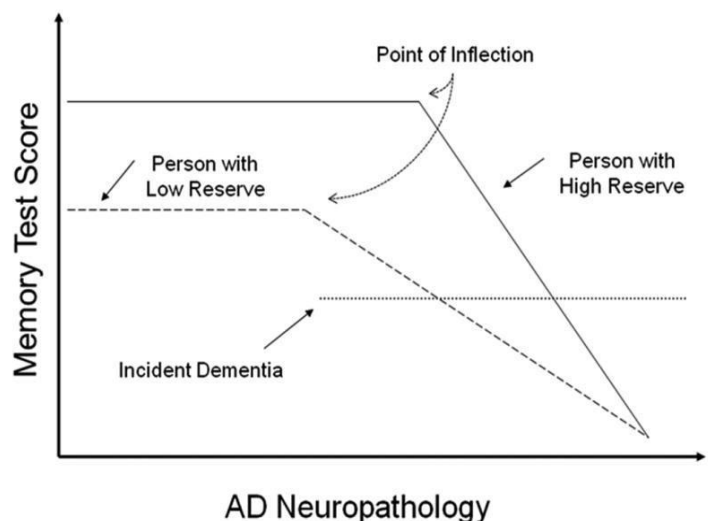


Figure 1. Theoretical illustration demonstrating the conceptualized memory performance trajectory of individuals with high and low CR levels as a function of AD neuropathology (Stern, 2015). Low reserve individuals start at a lower level of cognitive functioning than those of high reserve, but also experience performance decline at an earlier time of disease advancement. High reserve individuals

maintain higher cognitive performance and experience a decline in later stages of neuropathology, however they also suffer a sharper decline in functioning once this decline (point of inflection) occurs.

Possessing a higher CR is still preferable however from a quality of life and even cost-of-care standpoint. If the inflection point is delayed, the time of life without debilitating symptoms is increased and thus less overall outside care is needed. This is progressively becoming more important to understand given the rise in average population age (United States Census Bureau, 2020). Along with this increase in aged individuals follows a rise in age-related brain diseases as well, such as Alzheimer's, Parkinson's, and dementia (Sauer, 2018). These diseases are devastating to the individual's cognitive health, but carry extra weight given the financial cost associated with treatment. A paper published in the *Annals of Neurology* estimated that just nine of the most common neurological diseases cost Americans \$789 billion dollars (Gooch, C. L., 2017). That cost is projected to rise and eventually double by the year 2050.

These costs could potentially be mitigated if the public was more aware of how to delay disease symptomology and overall cognitive decline. Recent market data in fact indicate an increasing public interest in cognitive health. In 2018, almost \$2 billion were spent on phone applications aimed at keeping the brain sharp (Tov, 2019). Similarly, data collected from Barnes & Noble's website, a popular website for buying books, demonstrated that American readers have become much more interested in books about brain health than those involving diet and exercise (Schaub, 2019). If preventative measures against cognitive decline were more well-studied, the public could then be more aware of the factors in their lives to change, potentially improving the quality of their lives and decreasing the medical costs they would have accrued otherwise.

Some factors which influence cognitive decline, discovered through decades of CR research, include genetics (Ando, J., Ono, Y. & Wright, M.J., 2001), diet (Clare, L. et al., 2017), educational attainment (Staff, R.T. et al., 2004), occupation (Nucci et al., 2012), and lifestyle (Scarmeas et al., 2003), among others. While education and career level attainment is impractical to change late in life, and for genetics, impossible, the leisure activities an individual engages in during his free time are substantially more modifiable. The impact of these activities on the brain too are far from negligible. Helzner et al. (2007) found that the number of leisure activities engaged in by an individual correlates with cognitive measures in a very consistent way to the conceptualized CR trajectories of cognitive performance (Figure 1). That is to say, those who have engaged in relatively more leisure activities display a cognitive

trajectory comparable to the trajectory one would see in a high reserve individual. In fact, this association between leisure activity engagement and cognitive functioning has repeatedly been observed (Barnes, J. N., 2015; Pinto, C. & Tandel, K., 2016; Stern, Y., 2012).

Furthermore, a study using positron emitting tomography (PET) for the study of blood flow in AD patients, found that individuals with higher leisure activity rates could tolerate greater deficits in cerebral blood flow, even when controlling for IQ, educational level, and disease severity (Scarmeas, 2003). Scarmeas et al. (2001) also found that individuals who engaged in more than six leisure activities experienced a 38% lower risk of developing dementia than low leisure activity-engaging individuals. This research substantiates leisure activity as a significant component of CR and importantly, a component which can be easily modulated by the individual.

However, much of the current available research that explore lifestyle measures as a predictor of CR do so as a universal measurement of overall leisure activity engagement frequency and do not provide finer categorical dissemination on the range of leisure activities that exist (Wang, H. X. et al., 2013; Stern & Yaakov, 2006; Scarmeas, N., & Stern, Y., 2003). The portion of studies that do include leisure activity sub-sets, typically use broad categorical activity descriptors and focus on only one of those activity types per study. That is not to say though that these studies do not provide important data to the field of CR research. Isolating categories of leisure activities which impact brain health, even in a broad manner, brings us a degree closer to understanding how to group such activities.

Evans and colleagues (2018), for example, found that social engagement was associated with an enhancement of CR and overall cognitive functioning. Conversely, social isolation was found to be detrimental to brain health. Physical activity as well has long been associated with cognitive performance (Nithianantharajah, J., & Hannan, A. J., 2009; Studenski, S., et al., 2006; Wilkee, G. & Martella, D., 2018). Sheung-Tak Cheng (2016) reviewed the associations between physical activity and the risk of dementia and found a positive correlation between engagement in aerobic exercise and lack of age-related gray and white matter loss (suggesting increased BR). As well, he discovered that intellectual activities, activities which require higher-level cognitive processing, were associated with increased plasticity of neural circuits (increased CR). This latter finding is also supported by previous research which linked intellectual enrichment to cerebral efficiency (Sumowski et al., 2010).

Unfortunately though, broad categorization of leisure activities in this manner is not entirely helpful from a clinical

treatment perspective given the diversity of activities. Half an hour a day of running does not activate the same areas of the brain as playing chess at the same frequency, for example, and thus may benefit the brain in a distinct manner (Cheng, S. T., 2016). It is because of this diversity in activation, and consequential cognitive effect, that I chose to look at the impact leisure activities have on the brain at a finer categorical level.

This categorization was performed in a systematic manner such that leisure activities were grouped using a traditional factor-based analysis (FBA) via SPSS. This removes the need for arbitrary decisions about leisure typology when grouping. This analysis results in statistically distinct groups of activities with corresponding participant scores which can be used to compare the differences in structural and cognitive measures between groups. More about this process will be explained further on.

The cognitive measures included in this study are a range of tests measuring varying functions of the brain such as working memory, mental flexibility, language, processing speed, and so on. These measures allow us to obtain a multi-faceted cognitive profile for each participant which, when grouped by others based on their leisure activity measures, can be used to illuminate cognitive differences between those in different leisure activity type groups (see Methodology).

It was hypothesized that comparing scores across leisure type groups in this manner will result in at least one significant difference. This hypothesis derives from the longstanding neurological understanding that different activities engaged in by an individual results in different neural activation. Further, by means of synaptic plasticity, neural activation results in a strengthening of the neural connections associated with the engaged activity (Hughes Jr., 1958).

Along with cognitive measures, structural differences in the brain were predicted to exist between leisure type groups due to the same rationale. Differential neural activation leads to differential neural maintenance and sometimes even neuronal growth (Kwok V. et. al., 2011). It is due to the preservation and growth of brain structure elicited by neural activation that those with differing histories of leisure activity engagement would manifest differences in brain volume as well. These differences can be found using popular statistical tests designed for comparing structural magnetic resonance images (MRI) of the brain. Voxel-based morphometry (VBM) is designed to compare overall gray-matter volume differences (GMDs) between subject groups. Region of interest analyses (ROI) are performed on distinct structures within the brain using structural masks in

conjunction with subject brain data. Both of these tests were implemented in this design.

Significant VBM results were expected when comparing those with higher and lower final CRI-Leisure score (CRI-LTm) due to the established link between overall leisure activity engagement and greater gray matter volume (Gow et. al., 2012). Such a finding would confirm a passive CR model effect of leisure activity engagement on the brain. This difference was not expected to be found in comparisons between leisure type groups in a macro-focused VBM analysis. This is because I did not expect differences in leisure activity engagement by type to have predictability on overall GMDs, but to impact a sub-set of brain areas distinctly. ROI analyses were expected to illuminate significant differences between these groups on a finer structural scale, though no specific predictions were made in these comparisons due to the novelty of the grouping procedure. Instead, a data-driven statistical approach was taken to determine if this study's approach to grouping could effectively find group differences. ROI analyses using all available brain maps within the Harvard-Oxford cortical and sub-cortical atlases were performed in order to capture any and all significant structural differences.

II. METHODOLOGY

i. COGNITIVE PERFORMANCE MEASURES

One hundred and eight Dutch-speaking participants were recruited using advertisements posted around Nijmegen, The Netherlands. Those who had a history of psychiatric or neurological disorder, cognitive impairment or MRI contraindications were excluded from participating. Prior to scanning, participants completed a battery of cognitive tests administrated by a research assistant. Along with those, participants completed a written questionnaire which measured their education level attainment, work history, and leisure activity engagement.

Cognitive Reserve Index Questionnaire (CRIq)

The CRIq, created by researchers at the University of Padova in Italy, is an instrument for measuring CR by recording data regarding participant work history, educational history, and participation in various leisure activities. Education is recorded by years of education as well as vocational training. Work activity is broken up into different levels of required skill, ranging from low skilled manual work to a highly responsible or intellectual occupation. Leisure activity is divided into weekly, monthly, annually, and fixed frequency groups. The items within these groups are then used to record years of participation frequency in each item since adulthood (age 18).

In order for each activity item to be recorded for frequency, the participant must at least participate in that activity item at least 3 times per the time of the frequency group. Reading books is an activity in the *weekly* frequency group so the participant must participate in that activity 3 or more times per *week* to include that activity in the total CRI-LTm score. Similarly, going to the cinema is a *monthly* activity in the CRIq and therefore must be done at least 3 times in a *month* to add to the overall score. The time recorded for each activity is the number of years of participation rounded up to the nearest increment of 5. This means that 6 years of activity participation, for example, would be recorded as 10 years.

Education, work experience, and leisure time are then calculated into total scores, and those scores are aggregated into a total CRI score which represents overall CR. In this study, only the CRI-LTm scores were used to group participants, with the CRI work and education scores being regressed out, along with age, to avoid confounding effects (Figure 2). Once participants were grouped according to their CRI-LTm scores, behavioral test averages and gray matter averages were compared between groups to determine statistically significant differences.

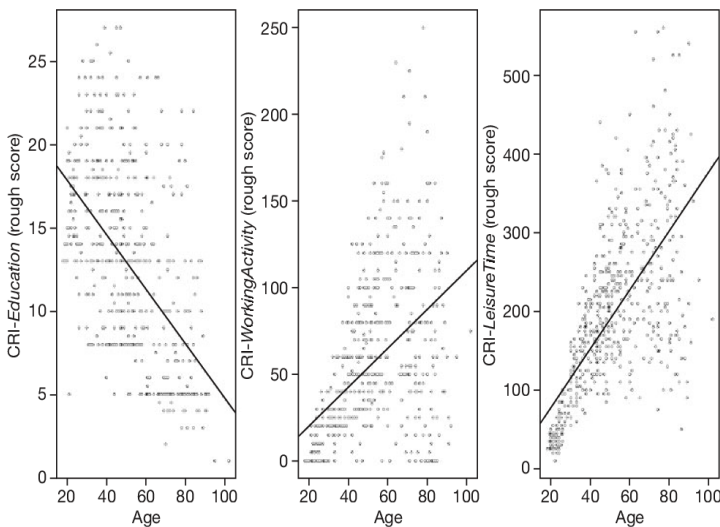


Figure 2. Scatterplots of raw CRI score values for education, work experience, and leisure time. Lines represent linear regression estimates, with scores as dependent variables and age as the independent variable (Nucci, 2012). All 3 CRI scores highly correlate with the age of the participant, thus age is regressed out of each.

The behavioral measures used tested a range of cognitive functions which gave a multi-faceted representation of an individual’s cognitive ability. Importantly, these tests measure distinct cognitive functions, allowing for more specific inferences to be made from my analyses. The included tests are the following:

🚩 Dutch Adult Reading Test (DART)

The DART consists of 50 individual words which vary in complexity. Simple words include *shock* and *fair*, while more complex words like *auxiliair* and *oecotype* are also included. The participant reads each word one-at-a-time while the test administrator marks their pronunciation as either *Good*, *Wrong*, or questionable (marked with a question mark).

The phonetic spellings of the words are provided to the test administrator, with the trickiest portion of the word underlined. The administrator then calculates the DART score by multiplying the total of *Good* ratings by 2, the total of in-between ratings by 1, and the total of *Bad* ratings by 0. This test has been found to accurately measure verbal intelligence and IQ (Schmand, B. et. al., 1991).

🚩 Montreal Cognitive Assessment (MoCA)

The MoCA is a cognitive assessment test which measures various cognitive functions, often used by healthcare professionals to detect cognitive impairments in patients. The MoCA test has been validated multiple times in its ability to distinguish normal controls from those with mild cognitive impairment or Alzheimer’s disease (Frietas, S. et. al., 2012).

Some of these test item include drawing a clock at “ten past eleven”, naming pictured animals, repeating back lists of digits, and naming words that begin with the letter F as quickly as possible within a minute. The items on the test explore aspects of cognitive functioning such as visuospatial/executive functioning, object recognition, memory, attention, language, abstraction, delayed recall, and time/place orientation (Nasreddine Z. S. et. al., 2005).

🚩 Stroop Color and Word Test (SCWT)

The SCWT is comprised of 3 sections. In the first section, participants are shown a card with 100 words of colors (red, green, yellow and blue), organized 10x10. The task is to read each word, from left to right, as quickly as possible. In the next section, participants are shown a card of colored strips and told to name the color of each strip as quickly as possible. Lastly, participants are shown a third card consisting of colored words, however the colors do not match the color name of the word. For example, the word *Blue* may be colored green, yellow, or red, but not blue. The task is to read the word while ignoring the incongruity of the color.

This mismatch of stimulus attributes creates a cognitive interference pattern known as the Stroop Effect (Scarpina,

F., & Tagini, S., 2017). The inhibition of this effect is thought to require cognitive flexibility, selective attention, and processing speed. All 3 sections are timed and the deviation of completion speed of the final word (W)/color (C) section compared to the completion times of the separate color and word sections is used as the participants SCWT score ($\frac{C\ score+W\ score}{2} - CW\ score = SCWT\ score$).

✚ Wechsler Memory Scale Revised (WMS-R)

The WMS-R is a neuropsychological test battery for assessing various memory functions in adults (Hunsley, J., & Lee, C. M., 2010). From the WMS-R, only the digit span test was used. The test consists of two sections of two trials each, with each trial consisting of six number sequences.

In the first section, a sequence of numbers is read to the participant, with the participant repeating each sequence afterwards. This begins on the 1st trial as a 3-digit sequence and can continue up to 8 digits. If the participant fails to accurately repeat a sequence, they continue immediately onto the 2nd trial, otherwise they complete all sequences and then continue onto the 2nd trial. During the 2nd trial, if the participant fails to report a digit sequence at the same length as the 1st trial (e.g., failing to accurately report a 5-digit sequence on both trials), the test is ended.

In the second section, participants perform a similar task, however they are asked to report the number sequences backwards. The digit sequence length begins at 2 digits and progresses up to 7. The WMS-R score is calculated as the total number of correctly repeated digit sequences.

✚ Trail Making Task (TMT)

The TMT is a popular neuropsychological test generally used to measure visual search speed, processing speed, and mental flexibility. It has also been found to be a sensitive measure for detecting Alzheimer's and other dementias (Cahn, D. A. et al., 1995). The TMT is now included in the Halstead-Reitan Neuropsychological Test Battery (HRNB) for use in assessing the condition and functioning of the brain (Darby, D. & Walsh, K.W., 2005; Walsh, K.W., 1991).

During the TMT, participants are presented with 25 numbered circles randomly placed on a sheet of paper. In part A, the task is to connect the 1st circle (circle "1") to circle "2" and so on, finishing at circle "25". In part B, participants must switch from increasing numeral value to increasing alphabetical order, and back again. Participants begin at circle "1" then draw to circle "A", then circle "2", then circle "B", and continue this pattern until circle "13". If the correct pattern is broken during either part, the

administrator will verbally correct them before they move onto the next circle. Both parts of the task are timed and the division of these times is used as the TMT score value. This final score value is then multiplied by -1 in order to standardize higher scores as indicators of better performance.

✚ Story Recall Task (SRT)

The SRT is another widely-used neuropsychological test, commonly administered to evaluate verbal memory functioning. It has been found to be reliable in distinguishing normal controls and those with mild cognitive impairment (MCI) or Alzheimer's disease (Baek M. J., et al., 2011). The version used in this study was taken from the RBMT-3 (Kurtz, M. M., 2011).

The task involves the test administrator reading a short passage of text to the participant while the participant is told to remember as many details as he can. The participant is then told to recount what he remembers while the administrator marks which aspects of the story were reported. Twenty-one specific details of the story are listed for the administrator to score.

The task consists of two parts: immediate recall and delayed recall. During the immediate recall section, the participant responds with remembered details as soon as the story is done being told and the participant is prompted to respond. During the delayed recall section, the participant must wait 20 minutes before responding. A perfectly remembered detail is worth 1 point, while a partially remembered detail is worth ½ point. The total SRT score is the sum of these points.

✚ Prospective and Retrospective Memory Questionnaire (PRMQ)

The PRMQ is a self-report questionnaire which measures prospective and retrospective memory. The questionnaire is comprised of 16 items which illustrate common memory errors most people experience, such as "*How often do you decide to do something in a few minutes and then forget?*" and "*How often do you not recognize a place where you have been before?*". Participants then rate the frequency with which they experience these errors using a 5-point scale from "*Very Often*" to "*Never*".

In order to calculate the final PRMQ score, each item is given a respective point value, with "*Never*" being worth 0 and "*Very Often*" worth 4. These values are then added together into an aggregate score. This aggregate score is then multiplied by -1.

ii. BEHAVIORAL ANALYSES

In order to validate the CRI leisure time measures as predictors of differences in cognitive performance and brain structure, an initial comparison of those with high versus low CRI-LTm scores was performed. The divide for this grouping was defined as those above and below the median CRI-LTm value. To diminish the effects of closeness in scores between the higher and lower tier groups, the bottom 12.5% and the top 12.5% were excluded from the higher and lower tier groups, respectively. Those with outlier data were also excluded to eliminate the possibility of biasing results.

Dimensionality reduction via FBA was performed using SPSS in order to separate leisure items into comparison groups in an unbiased manner using participants' CRI-LTm scores. However, to eliminate confounding effects of age, education, and work activity on leisure measurements, age, education, and work scores were regressed out of these scores beforehand. Eigenvalues were calculated for each leisure item grouping using an orthogonal rotation and the final number of comparison groups was determined by the number of total component eigenvalues greater than 1. The value of 1 was chosen because components with eigenvalues greater than 1 capture more variance than a single included variable. Extraction via principal component analysis revealed four unique groupings which pass this criteria (Figure 3). These four groupings explain 60.12% of the overall variance of all leisure items.

A

Component	Total	% of Variance	Cumulative %
1	6.528	38.402	38.402
2	1.366	8.035	46.436
3	1.203	7.078	53.514
4	1.123	6.608	60.122
5	0.991	5.828	65.950
...

B

Scree Plot of Leisure Activity Components

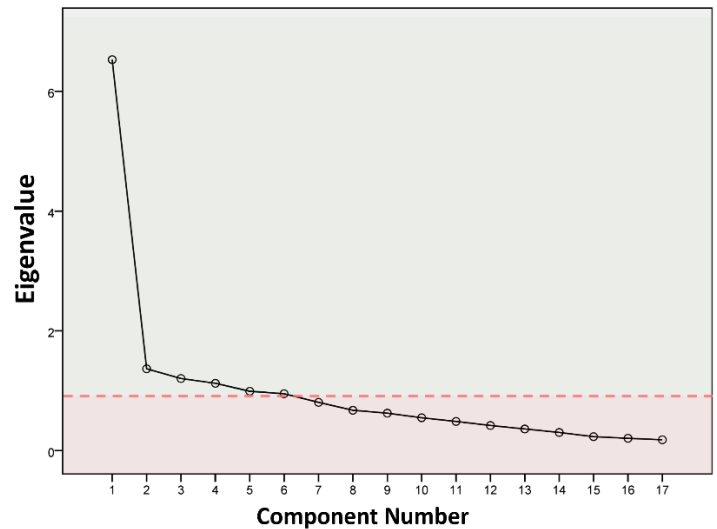


Figure 3. (A) Each component's (leisure activity grouping's) extracted eigenvalue and explained variance, and the total percentage of variance explained when including all above components. Eigenvalues were via principal component analysis using an orthogonal rotation. (B) Scree plot of the components' extracted eigenvalues. The red dotted line indicates the cut-off Eigenvalue of 1 which was used to determine the number of leisure activity groupings.

Once the number of groupings was determined, the loadings in the rotated component matrix were used in order to determine which group each leisure item belongs to. These values represent the correlation of each item to each group. Items were divided into groups simply using each item's highest group correlation value (Figure 4a).

Leisure Activity	1	2	3	4
Reading newspaper	.792	.267	.236	.025
Managing account	.767	.394	.096	.142
Using new tech	.755	.231	.239	.025
Raising children	.722	-.020	.172	.278
Driving	.705	.103	.312	.240
Doing chores	.619	.433	-.012	.450
Gardening	.473	.450	.007	-.126
Going to the movies	-.123	.712	.200	.161
Attending concerts	.329	.652	.025	.216
Reading books	.508	.564	.218	.049
Going on journeys	.228	.515	.111	.064
Art	.287	.361	.139	-.053
Socializing	.082	.155	.736	-.066
Volunteering	.236	.022	.697	.140
Leisure activities	.280	.269	.662	.126
Caring for grandchildren	.045	.008	.251	.861
Caring for pets	.304	.314	-.151	.617

Figure 4a. Rotated component matrix containing the factor loadings (also known as correlational values) of each leisure item. Each row contains a factor loading of a leisure item for each component grouping (column 1-4). The grouping of each leisure item was determined by the largest component factor loading for the item. The highest factor loading for newspaper, for example, is 0.792. Therefore newspaper is grouped in the 4th grouping.

Labels were then created for each grouping based on the leisure activities contained in each group (Figure 4b). The 1st group's activities primarily involve activities which are necessary for one to engage in, such as managing one's account, doing chores, and driving. This group is thus labeled *Responsibilities*. The 2nd group is comprised of activities which are only engaged in for enjoyment e.g., going to the movies, attending concerts, and going on journeys. Due to this, this group of activities is labeled *Entertainment*. The 3rd group includes activities which can be considered social (e.g., socializing), physically active (e.g., leisure activities), or both (e.g., volunteering). Therefore this group is given the label of *Social/Physical*. Lastly, the 4th group is labeled *Caretaking* due to both of the group's only activities involving caring for grandchildren or caring for pets.

Responsibilities	Entertainment	Social/Physical	Caretaking
Reading newspaper	Going to the movies	Socializing	Caring for grandchildren
Managing account	Attending concerts	Volunteering	Caring for pets
Using new technology	Reading books	Leisure activities	
Raising children	Going on journeys		
Driving	Art		
Doing chores			
Gardening			

Figure 4b. Leisure activities from the CRIQ organized by the FBA grouping. Labels for the groupings were created using the types of activities contained in each group.

Participants' CRI-LTm score average for each FBA leisure activity grouping was calculated and used to determine which group the participant belongs to, by their highest group average. These grouped participants were then compared against those in every other group on their cognitive performance scores. These scores were analyzed via a between-subjects analysis of variance (ANOVA) using a series of MATLAB scripts, which can be found in a GitHub repository via the following link: www.github.com/PsychMike/CRIq_Analysis. Each score comparison is calculated for significance using a p-value threshold of <0.05.

iii. MRI ACQUISITION

Participants' brains were scanned using a 32-channel head coil in a Siemens 3T Prisma MRI scanner at a voxel resolution of 2mm. Participants remained in the scanner for less than an hour while T1-, T2-, and T2*-weighted structural scans were acquired in the transverse plane, with 66 slices per volume.

iv. MRI ANALYSES

Once MRI data had been acquired, the T1-weighted images were then pre-processed using FSL's brain-processing function. These processes use an experimental design template to extract the appropriate MRI data, segment them into gray matter (GM), white matter (WM), and cortical spinal fluid (CSF), non-linearly register the GM images to the design template, and then concatenate those images into a singular 4D image.

 Voxel-based grey matter morphometry (VBM)

VBM is a widely-used statistical method for comparing high-resolution MRI data on a voxel-wise scale. The VBM method has been used consistently over the past two decades to analytically compare brain tissue concentrations between subject groups in an automated and unbiased fashion (Douaud, G. et al., 2007; Foubert-Samier, A. et al., 2012; Good, C.D. et al., 2001; Pergher, P. et al., 2020).

After an averaged MRI image was created for every comparison group, a VBM analysis was performed on the data using FSL’s *randomise* and *fslmaths* functions. This method involves performing a series of statistical tests on voxel-wise GM comparisons by means of linear regression modeling (Ashburner, J., & Friston, K. J., 2000).

VBM analyses were performed in this experiment for the purpose of comparing GM densities between group datasets. VBM results were output as a range of p-values (0.00-1.00) as well as displayed as a statistical map overlaid on FSL’s standard MNI template. The statistical mappings were created using cluster-based thresholding with a 2mm smoothing.

✚ Region of interest analysis (ROI)

ROI analyses provide statistical brain volume comparisons in a more specific manner than when using VBM. ROI analyses use templates of T1-weighted MRI images which are segmented into various cortical and subcortical structural areas. These segmentations, or masks, are then individually combined with experimental 4D data templates to distinctly run statistical testing on only those brain segments (Poldrack, R. A., 2007).

For this experiment, all of the 69 brain structures in the Harvard-Oxford cortical and subcortical structural atlases were used to run ROI analyses (Desikan, R. S. et al., 2006; Frazier, J. A. et al., 2005; Goldstein, J. M. et al., 2006). These atlases contain T1-weighted images of 37 health individuals aged 18 to 50. All images were individually segmented in a semi-automated fashion and then processed further using FSL functions.

III. RESULTS

a. BEHAVIORAL RESULTS

A between-subjects ANOVA comparing upper and lower tier LTm groups across behavioral measures revealed statistical significance in SRT ($p=0.04$), TMT ($p=0.04$), WMSR ($p=0.02$), SCWT ($p<0.04$), and DART ($p=0.04$) scores. The upper tier LTm group exhibited averaged scores higher than the lower tier group in all measures, though not all were significant (Figure 5).

Upper/Lower Tier LTm Groups Across Behavioral Tests

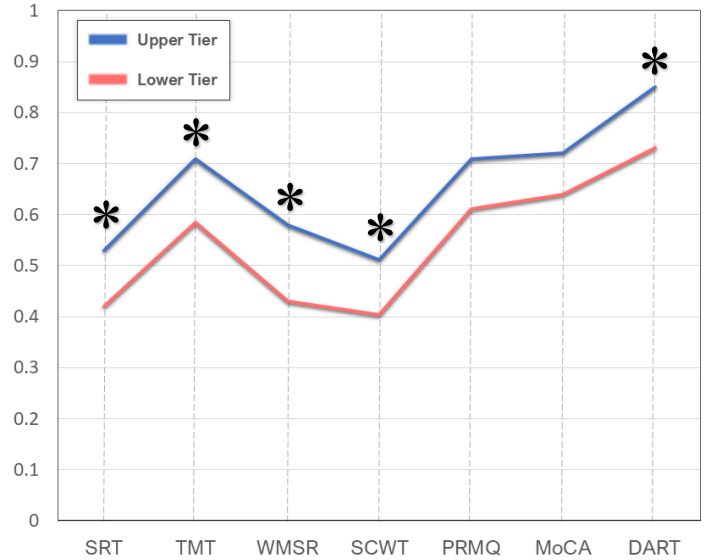


Figure 5a. A comparison of the upper (n=26) and lower (n=26) tier LTm groups across group averages of all behavioral measures using a normalized scale. The asterisks indicate significant ANOVA results for those comparisons.

SRT	TMT	WMSR	SCWT	PRMQ	MoCA	DART
0.038	0.046	0.020	0.045	0.272	0.333	0.042

Figure 5b. P-values for each behavioral test compared between upper and lower tier LTm groups using a between-subjects ANOVA. Significant p-values (<0.05) are highlighted in orange.

Comparing LTm scores of all the FBA sub-groups against each other revealed that the *Responsibilities* group achieved significantly higher TMT ($p=0.04$) and SCWT ($p=0.03$) scores than the *Entertainment* group. Also, the *Social/Physical* group achieved higher SRT ($p=0.02$) and SCWT scores ($p=0.01$) than the *Caretaking* group. All other comparisons did not reach statistical significance (Figure 6).

Groups compared	SRT	TMT	WMSR	SCWT	PRMQ	MoCA	DART
Responsibilities - Entertainment	0.701	0.045	0.919	0.025	0.118	0.690	0.601
Responsibilities - Social/Physical	0.641	0.181	0.575	0.396	0.487	0.843	1
Responsibilities - Caretaking	0.353	0.667	0.198	0.093	0.367	0.217	0.182
Entertainment - Social/Physical	0.449	0.789	0.249	0.877	0.848	0.788	0.738
Responsibilities - Caretaking	0.859	0.858	0.285	0.371	0.153	0.302	0.347
Social/Active - Caretaking	0.021	0.252	0.769	0.014	0.677	0.407	0.376

Figure 6. P-values for every group comparison via between-subjects ANOVA. Significant p-values (<0.05) are highlighted in orange.

b. MRI RESULTS

i. Upper vs. Lower Tier LTm

Eleven participants in the upper tier LTm group and eleven from the lower tier group were used in the MRI analyses due to their MR data being available. Due to various participants only completing the behavioral tests, compounded with the COVID-19 pandemic interfering with in-person testing, only a sub-set of participants from the behavioral analyses could be compared via MRI analyses.

A VBM analysis of those in the upper and lower tier LTm groups revealed no significant gray matter volume differences (GMDs) (min $p=0.61$). Multiple modifications of the MR templating and segmentation were performed, which resulted in similar non-significant results. As well, ROI analyses resulted in no significant results using any of the included brain masks.

ii. FBA Group Comparisons

Comparing FBA groups also resulted in no significant VBM results. However, ROI analyses revealed multiple significant GMDs in various brain areas. The *Responsibilities* group demonstrated higher GMDs than the *Social/Physical* in the lateral occipital cortex ($p<0.001$) and the occipital pole ($p=0.01$) (Figure 7).

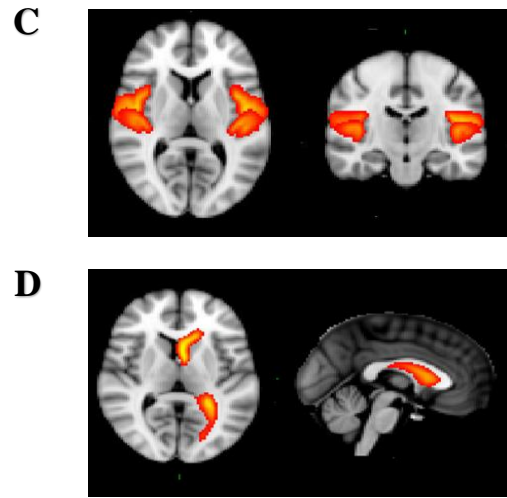
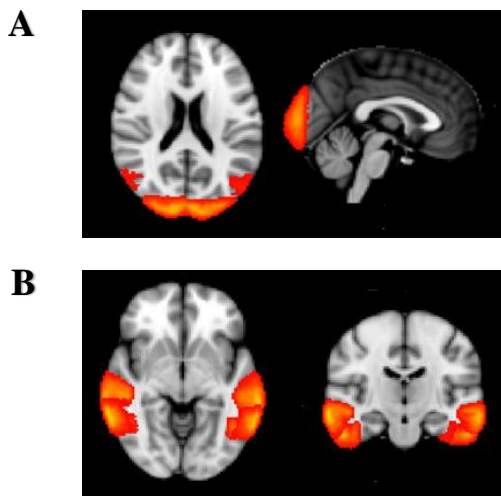


Figure 7. Areas of significant GMDs across groups, controlled for multiple comparisons via Bonferroni correction. (A) *Responsibility* > *Entertainment* (group $n=10$): lateral occipital cortex and the occipital pole. (B) *Responsibility* > *Social/Physical* (group $n=11$): inferior temporal gyrus, mid-temporal gyrus, and the fusiform gyrus. (C) *Responsibility* > *Caretaking* (group $n=10$): central opercular cortex and the Heschl's gyrus. (D) *Entertainment* > *Caretaking* (group $n=10$): left lateral ventricle.

The *Responsibilities* group also demonstrated higher GMDs than the *Social/Physical* group in the inferior temporal gyrus ($p=0.04$), the mid-temporal gyrus ($p=0.04$), and the fusiform gyrus ($p=0.04$), and higher GMDs than the *Caretaking* group in the central opercular cortex ($p=0.01$), the Heschl's gyrus ($p=0.01$), and the parietal operculum ($p=0.0$). Lastly, the *Entertainment* group demonstrated higher GMDs in the left lateral ventricle than the *Caretaking* group ($p=0.02$).

IV. EXPLORATORY ANALYSES

a. Upper vs Lower Tier Cognitive Effort

Another method of grouping leisure activities to be explored is by using the cognitive effort associated with each activity. If more cognitive effort is required to participate in the leisure activity, that may translate to a more substantial cognitive benefit.

To test this, a Google Forms survey was created which listed each CRI-LTm item with the exact item descriptions provided in the CRIq. One hundred random individuals online completed the survey by ranking each leisure item by the amount of cognitive effort these activities may require using a 1 (easy) to 5 (difficult) scale.

These subjective rankings were then averaged for each item and bins of upper and lower tier cognitive effort groups were created using an almost-even split (nine LTm items in the

upper tier, eight in the lower tier). From highest-to-lowest average ranking, the upper tier group consisted of *children, grandchildren, journeys, art, driving, reading books, social activities, using new technology, managing one's account* and the lower tier group consisted of *concerts, volunteer work, general leisure activities, pets, reading the newspaper, gardening, doing chores, and going to the cinema*.

communication between members participating in the activity. Activities which do not meet this criteria are designated as *cognitive* activities (CA). By systematically reviewing MRI studies which feature either or both of these activity types, she discovered that both are associated with greater white matter density and gray matter volume.

However, while the reviewed studies included activities that could be labeled as *cognitive* or *social*, the authors of those studies did not make the same type-distinction within their study designs. Therefore, Anatürk's review relied on comparing the effects of cognitive and social activities from varying study designs and analytical parameters. A within-study implementation of his type-distinction may be useful to ensure that the activities used in the analyses closely follow her distinction and use exactly the same analytical procedures.

By using his criteria in this study, CA activities were determined to include *reading the newspaper, doing chores, driving, using new technology, gardening, art, reading books, and account maintenance*. SA activities included *social activities, caring for grandchildren, and caring for children*. Activities such as *general leisure activities, going on journeys, going to the cinema, and volunteer work* were excluded from either grouping due to the ambiguity of whether or not there is direct social interaction involved in engaging in them.

Participants were grouped by whether their CA or SA activity scores were higher and then compared with behavioral and MRI analyses. A between-subjects ANOVA comparing CA and SA groups revealed statistical significance only in DART scores ($p=0.01$) (Figure 9).

Upper/Lower Tier Effort Groups Across Behavioral Tests

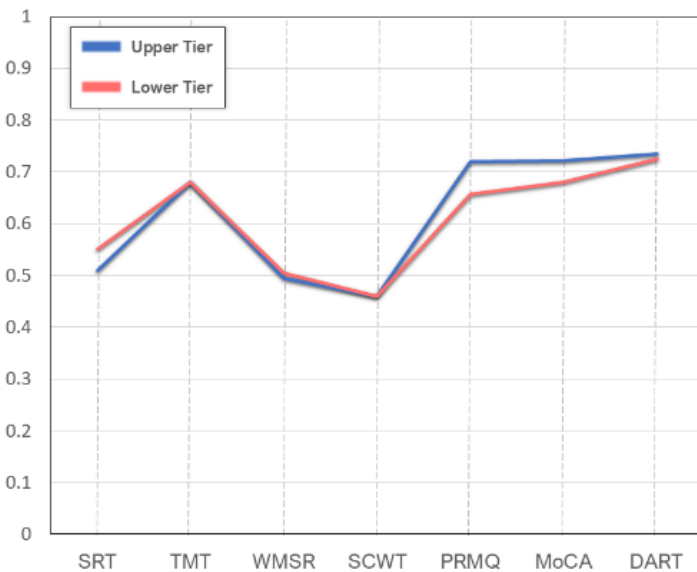


Figure 8a. A comparison of upper ($n=21$) and lower ($n=21$) tier LTm effort groups across group averages of all behavioral measures.

SRT	TMT	WMSR	SCWT	PRMQ	MoCA	DART
0.815	0.831	0.830	0.938	0.225	0.392	0.602

Figure 8b. P-values from comparing upper and lower tier LTm effort groups via between-subjects ANOVA.

A between-subjects ANOVA comparing these upper and lower tier LTm effort groups revealed no statistical difference between these groups on any behavioral measure (Figure 8). Comparing MR images, VBM analysis revealed no statistical GMDs between upper and lower effort groups ($\min p=0.10$). Also, no ROI analyses revealed GMDs in any included brain area.

b. Cognitive vs. Social Activities

Another proposed method of grouping leisure activities is by determining whether each activity is cognitively- or socially-oriented. Anatürk (2018) made this distinction using a criteria labeling social activities (SA) as those involving the presence of others and the explicit

Cognitive/Social Groups Across Behavioral Tests

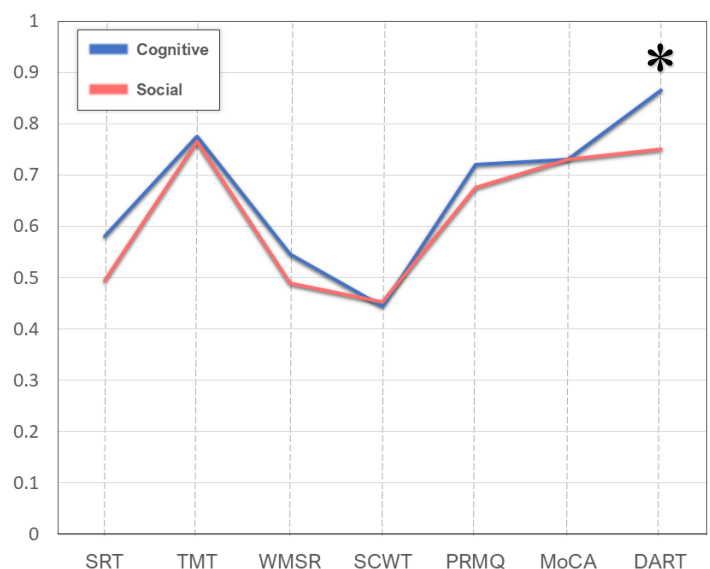


Figure 9a. A comparison of CA (n=28) and SA (n=28) groups across group averages of all behavioral measures.

SRT	TMT	WMSR	SCWT	PRMQ	MoCA	DART
0.072	0.082	0.634	0.941	0.903	0.381	0.013

Figure 9b. P-values comparing social and cognitive activity groups across all behavioral measures via between-subjects ANOVA. Significant p-values (<0.05) are highlighted in orange.

Despite finding a behavioral difference between CA and SA groups, VBM analysis revealed no statistical difference between them (min. $p=0.66$). Also, significant GMDs were not found in any of the performed ROI analyses.

V. DISCUSSION

The goal of this study was to determine whether differential effects on cognitive and MRI measures could be found when dividing leisure activities into groups by activity type. This was done using FBA on leisure data collected from 108 CRI questionnaires. From this analysis, leisure items were divided into four statistically distinct groups, and participant data were then divided into these groupings. These groups' behavioral data were compared via ANOVA tests and their MRI data were compared via VBM and ROI analyses. Data were grouped in additional analyses by upper and lower tier CRI-LTm scores, by upper and lower tier scores ranked by cognitive effort, and by binning data into either *cognitive* or *social* activity groups.

Counter to my initial hypothesis, the comparison of upper and lower tier LTm groups revealed significant cognitive performance differences on various behavioral measures, but no significant GMDs via VBM or various ROI analyses. This seems to suggest a potential effect of leisure activity engagement on cognition, but not in a manner which impacts gray matter density. This would suggest a CR component to the CRI-LTm activities' impact on the brain in which upper tier LTm-grouped participants more efficiently utilize available cognitive resources with similar brain density to lower tier LTm participants. This finding consequently more closely matches the active model of CR than the predicted passive model.

Analyses performed on participants FBA-grouped by LTm categories demonstrated significant cognitive performance group differences in two behavioral test comparisons and significant GMDs in four ROI analyses. CR components may play a role in these analyses as well since even the FBA group comparison which resulted in both behavioral and structural differences, did not evince clear functional

connections between the two. With a finding that the *Responsibilities* group demonstrated higher TMT (spatial processing) and SCWT (mental flexibility) scores than *Entertainment* group, one may expect brain areas which are involved in performing those tests to exhibit relatively greater density as well, such as the posterior parietal cortex and the prefrontal cortex. Instead, only visual processing areas such as the lateral occipital cortex and the occipital pole were found to have greater density. While visual processing plays a role in completing these behavioral tasks, the function is secondary to higher level processes which the tests are designed to measure. The other group comparison which resulted in significant behavioral score differences, *Social-Physical – Caretaking*, demonstrated no significant GMDs in the VBM or ROI analyses. As with the upper and lower tier LTm comparisons, the explanation for this disparity may be that the brain areas recruited during the behavioral testing are simply functioning more efficiently in one group over the other and therefore greater density is not required to elicit higher cognitive performance.

In the exploratory analyses, grouping participant data by cognitive effort or using a CA/SA split revealed no significant GMDs. As well, no cognitive performance differences were found when comparing higher and lower tier cognitive effort groups, and only one behavioral measure (DART) significantly differed between CA and SA groups. The significantly higher DART (reading test) result may be due to the items, *reading a book* and *reading the newspaper* belonging to the CA group, and thus is not very surprising.

Anatürk found engagement in activities which combine cognitive and social elements, termed *social-intellectual* activities, to be indicative of greater white and gray matter density. An attempt to parse these elements by grouping activities as solely cognitive or social resulted in findings that indicate no behaviorally or structurally measurable difference between the two. This suggests that these components need not be categorically separated when grouping activities. Though, to further confirm this conclusion, upper and lower tier splits of CA and SA can be performed to test predictability of each on behavioral and MRI measures. Due to time constraints, these analyses were not performed within the current study.

Further exploring the findings of the FBA grouping approach, many of the results seem difficult to conceptually resolve. This challenge includes interpreting seeming unrelated behavioral and gray matter differences within the same group comparison, as well as why particular groups differ significantly on specific cognitive measures which appear irrelevant to the groups' activities. For example, the ROI results suggest a statistical GMD between the

Responsibilities group and the *Entertainment* group in the lateral occipital cortex and the occipital pole, both areas involved in visual processing, but since both group's activities involve vision, the reasoning for this is unclear. Moreover, the *Responsibilities* group demonstrated higher GMDs than the *Social/Physical* group in the temporal gyrus (auditory processing) and the fusiform gyrus (facial recognition) but auditory processing does not seem specific to either group's activities. Counterintuitively, facial processing areas were found to be smaller in the *Social/Physical* group, a group of activities which includes *socializing*, an activity which logically involves processing the faces of others.

Perhaps easier to interpret, comparing the *Responsibilities* group to the *Caretaking* group, the *Responsibilities* group exhibited higher GMDs in the operculum (thought, cognition, and planning) and the Heschl's gyrus (language processing). Cognitive benefits of the *Responsibilities* group's activities, compared to the *Caretaking* group's, could be argued due to the required need for thought and planning in many of the group's activities, such as raising children, doing chores, and gardening. However, it is not directly apparent why results indicate higher TMT (spatial processing) and SCWT (mental flexibility) scores in the *Responsibilities* group than the *Entertainment* group, nor why the *Social/Physical* group exhibited higher SRT (verbal memory) and SCWT scores than the *Caretaking* group.

Truthfully, these results may not necessarily need to be immediately explainable so long as they hold statistical significance. That is to say, as long as the activities are predictive of better cognitive performance and higher gray matter density when controlling for confounding variables, finding a logical connection between each activity and the cognitive processes the tests purport to measure is a secondary concern saved for future research. From a clinical treatment perspective, the confirmation of these causal relationships alone are much more important than why they exist.

That aside, it does seem imprudent to forgo interpreting the finding that the *Responsibilities* group demonstrated higher GMDs than all other FBA groups. A potential explanation for this may be because more engagement in responsibilities is indicative of the individual's automaticity and ability, cognitively, to regularly take care of those responsibilities. Furthermore, since responsibilities are mandatory activities rather than those engaged in simply for enjoyment, higher responsibility engagement may be due to a more structured lifestyle and possibly a higher level of maturity. The latter of which, at least, has been found to be correlative to intelligence (Cohn, L. D., & Westenberg, P. M., 2004; Landau, E., & Weissler, K., 1998). This would also explain

why greater cognitive differences are found for the *Responsibilities* group, a group of lifestyle activities which must be completed, over the *Entertainment* group, a group of optional activities which fit conventional definitions of leisure. This, however, is purely speculative and not revealed in this study through corollary analyses.

Overall, it is encouraging that FBA grouping of leisure activity data resulted in multiple behavioral and structural group differences. This approach provides a statistically sound and time-efficient manner in which to study leisure activities at the level of category instead of exclusively by general activity engagement. Using standardized parameters to FBA-process leisure activity data results in grouping which is much less confounded with subjective decision-making usually involved in activity grouping, such as developing arbitrary definitions of category types or by using survey data. This can later be further explored beyond the scope, and limitations, of this study.

Due to the constraints of collecting new data during the COVID-19 situation, only 108 participants' data were included in this study's analyses. In future studies, it might be worthwhile to incorporate a database of many more participants in order to allow for more strict leisure activity type groupings without losing significant statistical power. With a very large subject database, for example, one could analyze the effects of leisure activities individually in place of, or in conjunction with, FBA grouping. This could potentially reveal specific leisure activities which more strongly benefit the individual than other activities within their grouping. With the current FBA grouping, it is plausible that a positive effect of engagement on the brain in one activity is offset by a negative effect of another activity within the same group. Distinct analyses on individual leisure activities' effects would reveal this.

It may be also beneficial to have a more extensive leisure activity questionnaire that could incorporate a broader range of activities with more specificity. The CRI-LTm item, *leisure activities*, for instance, could logically be broken down further into more specific activity items. This could simply be done by including numerous additional leisure activities on the given questionnaire and then FBA grouping them to find redundancy between activity engagement.

VI. CONCLUSION

The current study adds to increasing CR research which suggest that grouping leisure activities is imperative to studying the differential effects of activity engagement on the brain. These effects become obscured when using only broad leisure activity measures, and therefore stricter categorical grouping methods are needed. The FBA

approach provides a fast and relatively objective manner in which to create these groupings. The results of this study suggest grouping leisure activities using this technique is effective at finding both behavioral and structural differences between leisure activity types and substantiates the need to do so. Future directions for research would include expanding the pool of included leisure activities and measuring the effects of those activities individually in order to improve grouping.

REFERENCES

- [1] Anattürk, M., Demnitz, N., Ebmeier, K. P., & Sexton, C. E. (2018). A systematic review and meta-analysis of structural magnetic resonance imaging studies investigating cognitive and social activity levels in older adults. *Neuroscience and biobehavioral reviews*, *93*, 71–84. <https://doi.org/10.1016/j.neubiorev.2018.06.012>
- [2] Andersson J. L. R., Jenkinson M., Smith S. (2010). Non-linear registration, aka spatial normalisation. FMRIB technical report TR07JA2 Ando, J., Ono, Y., & Wright, M. J. (2001). Genetic structure of spatial and verbal working memory. *Behavior genetics*, *31*(6), 615–624. <https://doi.org/10.1023/a:1013353613591>
- [3] Ando, J., Ono, Y., & Wright, M. J. (2001). Genetic structure of spatial and verbal working memory. *Behavior genetics*, *31*(6), 615–624. <https://doi.org/10.1023/a:1013353613591>
- [4] Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry--the methods. *NeuroImage*, *11*(6 Pt 1), 805–821. <https://doi.org/10.1006/nimg.2000.0582>
- [5] Baek, M. J., Kim, H. J., Ryu, H. J., Lee, S. H., Han, S. H., Na, H. R., Chang, Y., Chey, J. Y., & Kim, S. (2011). The usefulness of the story recall test in patients with mild cognitive impairment and Alzheimer's disease. *Neuropsychology, development, and cognition. Section B, Aging, neuropsychology and cognition*, *18*(2), 214–229. <https://doi.org/10.1080/13825585.2010.530221>
- [6] Barnes J. N. (2015). Exercise, cognitive function, and aging. *Advances in physiology education*, *39*(2), 55–62. <https://doi.org/10.1152/advan.00101.2014>
- [7] Cahn, D. A., Salmon, D. P., Butters, N., Wiederholt, W. C., Corey-Bloom, J., Edelstein, S. L., & Barrett-Connor, E. (1995). Detection of dementia of the Alzheimer type in a population-based sample: neuropsychological test performance. *Journal of the International Neuropsychological Society : JINS*, *1*(3), 252–260. <https://doi.org/10.1017/s1355617700000242>
- [8] Cheng, S. T. (2016). Cognitive Reserve and the Prevention of Dementia: the Role of Physical and Cognitive Activities. *Current psychiatry reports*, *18*(9), 85. <https://doi.org/10.1007/s11920-016-0721-2>
- [9] Clare, L., Wu, Y. T., Teale, J. C., MacLeod, C., Matthews, F., Brayne, C., Woods, B., & CFAS-Wales study team (2017). Potentially modifiable lifestyle factors, cognitive reserve, and cognitive function in later life: A cross-sectional study. *PLoS medicine*, *14*(3), e1002259. <https://doi.org/10.1371/journal.pmed.1002259>
- [10] Cohn, L. D., Westenberg, P. M., & Cohn, L. D. (2004). Intelligence and maturity: meta-analytic evidence for the incremental and discriminant validity of Loevinger's measure of ego development. *Journal of personality and social psychology*, *86*(5), 760–772. <https://doi.org/10.1037/0022-3514.86.5.760>
- [11] Darby, D., & Walsh, K.W. (2005). *Walsh's Neuropsychology: A Clinical Approach*, (5th ed.). Edinburgh: Elsevier/Churchill Livingstone.
- [12] Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., Dale, A. M., Maguire, R. P., Hyman, B. T., Albert, M. S., & Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. <https://doi.org/10.1016/j.neuroimage.2006.01.021>
- [13] Douaud, G., Smith, S., Jenkinson, M., Behrens, T., Johansen-Berg, H., Vickers, J., James, S., Voets, N., Watkins, K., Matthews, P. M., & James, A. (2007). Anatomically related grey and white matter abnormalities in adolescent-onset schizophrenia. *Brain : a journal of neurology*, *130*(Pt 9), 2375–2386. <https://doi.org/10.1093/brain/awm184>
- [14] Evans I. E. M., Llewellyn D. J., Matthews F. E., Woods R. T., Brayne C., Clare L. et al. (2018). Social isolation, cognitive reserve, and cognition in healthy older people. *PLoS ONE* *13*(8): e0201008. <https://doi.org/10.1371/journal.pone.0201008>
- [15] Foubert-Samier, A., Catheline, G., Amieva, H., Dilharreguy, B., Helmer, C., Allard, M., & Dartigues, J. (2012). Education, occupation, leisure activities, and brain reserve: A population-based study. *Neurobiology of Aging*, *33*(2). <https://doi.org/10.1016/j.neurobiolaging.2010.09.023>
- [16] Frazier, J. A., Chiu, S., Breeze, J. L., Makris, N., Lange, N., Kennedy, D. N., Herbert, M. R., Bent, E. K., Koneru, V. K., Dieterich, M. E., Hodge, S. M., Rauch, S. L., Grant, P. E., Cohen, B. M., Seidman, L. J., Caviness, V. S., & Biederman, J. (2005). Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder. *The American journal of psychiatry*, *162*(7), 1256–1265. <https://doi.org/10.1176/appi.ajp.162.7.1256>
- [17] Freitas, S., Simões, M. R., Alves, L., Vicente, M., & Santana, I. (2012). Montreal Cognitive Assessment (MoCA): validation study for vascular dementia. *Journal of the International Neuropsychological Society : JINS*, *18*(6), 1031–1040. <https://doi.org/10.1017/S135561771200077X>
- [18] FSL Atlases. (n.d.). Retrieved from <https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases>
- [19] Goldstein, J. M., Seidman, L. J., Makris, N., Ahern, T., O'Brien, L. M., Caviness, V. S., Jr, Kennedy, D. N., Faraone, S. V., & Tsuang, M. T. (2007). Hypothalamic abnormalities in schizophrenia: sex effects and genetic vulnerability. *Biological psychiatry*, *61*(8), 935–945. <https://doi.org/10.1016/j.biopsych.2006.06.027>
- [20] Gooch, C. L., Pracht, E., & Borenstein, A. R. (2017). The burden of neurological disease in the United States: A summary report and call to action. *Annals of neurology*, *81*(4), 479–484. <https://doi.org/10.1002/ana.24897>
- [21] Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Friston, K. J., & Frackowiak, R. S. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*, *14*(1 Pt 1), 21–36. <https://doi.org/10.1006/nimg.2001.0786>
- [22] Gow, A. J., Bastin, M. E., Maniega, S. M., Hernandez, M. C., Morris, Z., Murray, C., . . . Wardlaw, J. M. (2012). Neuroprotective lifestyles and the aging brain: Activity, atrophy, and white matter integrity. *Neurology*, *79*(17), 1802–1808. <https://doi.org/10.1212/wnl.0b013e3182703fd2>
- [23] Helzner, E. P., Scarmeas, N., Cosentino, S., Portet, F., & Stern, Y. (2007). Leisure activity and cognitive decline in incident Alzheimer disease. *Archives of neurology*, *64*(12), 1749–1754. <https://doi.org/10.1001/archneur.64.12.1749>
- [24] Hughes Jr. (1958). Post-tetanic potentiation. *Physiological reviews*, *38*(7), 91–113. <https://doi.org/10.1152/physrev.1958.38.1.91>
- [25] Hunsley, J., & Lee, C. M. (2010). *Introduction to clinical psychology: An evidenced-based approach*. John Wiley & Sons Inc.
- [26] Institute for Digital Research & Education (n.d.) *A Practical Introduction To Factor Analysis: Exploratory Factor Analysis*. Retrieved from <https://stats.idre.ucla.edu/spss/seminars/introduction-to-factor-analysis/a-practical-introduction-to-factor-analysis/>
- [27] J. Ashburner, K. Friston (2000). Voxel-based morphometry — the methods. *Neuroimage*, *11* (6), pp. 805–821.
- [28] Jenkinson, M., Pechaud, M., and Smith, S (2005). BET2: MR-based estimation of brain, skull and scalp surfaces. *Eleventh Annual Meeting of the Organization for Human Brain Mapping*.
- [29] Katzman, R., Terry, R., De Teresa, R. et al. (1988). Clinical, Pathological, and Neurochemical Changes in Dementia: A Subgroup with Preserved Mental Status and Numerous Neocortical Plaques. *Annals of Neurology*, *23*, 138–144. <http://dx.doi.org/10.1002/ana.410230206>
- [30] Kurtz M. M. (2011). Rivermead Behavioral Memory Test. In: Kreutzer J.S., DeLuca J., Caplan B. (eds) *Encyclopedia of Clinical Neuropsychology*. Springer, New York, NY. https://doi.org/10.1007/978-0-387-79948-3_1154
- [31] Kwok, V., Niu, Z., Kay, P., Zhou, K., Mo, L., Jin, Z., . . . Tan, L. H. (2011). Learning new color names produces rapid increase in gray matter in the intact adult human cortex. *Proceedings of the National Academy of Sciences*, *108*(16), 6686–6688. <https://doi.org/10.1073/pnas.1103217108>
- [32] Landau, E., & Weissler, K. (1998). The Relationship between Emotional Maturity, Intelligence and Creativity in Gifted Children. *Gifted Education International*, *13*(2), 100–105. <https://doi.org/10.1177/026142949801300202>
- [33] Makris, N., Goldstein, J. M., Kennedy, D., Hodge, S. M., Caviness, V. S., Faraone, S. V., Tsuang, M. T., & Seidman, L. J. (2006). Decreased volume of left and total anterior insular lobe in schizophrenia. *Schizophrenia research*, *83*(2-3), 155–171. <https://doi.org/10.1016/j.schres.2005.11.020>
- [34] Mandolesi, L., De Bartolo, P., Foti, F., Gelfo, F., Federico, F., Leggio, M. G., & Petrosini, L. (2008). Environmental enrichment provides a cognitive reserve to be spent in the case of brain lesion.

- Journal of Alzheimer's disease* : JAD, 15(1), 11–28. <https://doi.org/10.3233/jad-2008-15102>
- [35] Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. <https://doi.org/10.1111/j.1532-5415.2005.53221.x>
- [36] Nithianantharajah, J., & Hannan, A. J. (2009). The neurobiology of brain and cognitive reserve: mental and physical activity as modulators of brain disorders. *Progress in neurobiology*, 89(4), 369–382. <https://doi.org/10.1016/j.pneurobio.2009.10.001>
- [37] Nucci, M., Mapelli, D., & Mondini, S. (2012). Cognitive Reserve Index questionnaire (CRIq): a new instrument for measuring cognitive reserve. *Aging clinical and experimental research*, 24(3), 218–226. <https://doi.org/10.3275/7800>
- [38] Pergher, V., Demaerel, P., Soenen, O., Saarela, C., Tournoy, J., Schoenmakers, B., Karrasch, M., & Van Hulle, M. M. (2019). Identifying brain changes related to cognitive aging using VBM and visual rating scales. *NeuroImage. Clinical*, 22, 101697. <https://doi.org/10.1016/j.nicl.2019.101697>
- [39] Pergher, V., Schoenmakers, B., Demaerel, P., Tournoy, J., & Van Hulle, M. M. (2020). Differential Impact of Cognitive Impairment in MCI Patients: A Case-Based Report. *Case reports in neurology*, 12(2), 222–231. <https://doi.org/10.1159/000507977>
- [40] Pinto, C. & Tandel, K. (2016). Cognitive reserve: Concept, determinants, and promotion. *Journal of Geriatric Mental Health*. 3 (44). <https://doi.org/10.4103/2348-9995.181916>
- [41] Poldrack R. A. (2007). Region of interest analysis for fMRI. *Social cognitive and affective neuroscience*, 2(1), 67–70. <https://doi.org/10.1093/scan/nsm006>
- [42] Sauer (2018). *Alzheimer's Is on the Rise in These States*. Retrieved from <https://www.alzheimers.net/alzheimers-is-on-the-rise-in-these-states>
- [43] Scarmeas, N., & Stern, Y. (2003). Cognitive reserve and lifestyle. *Journal of clinical and experimental neuropsychology*, 25(5), 625–633. <https://doi.org/10.1076/jcen.25.5.625.14576>
- [44] Scarmeas, N., Levy, G., Tang, M. X., Manly, J., & Stern, Y. (2001). Influence of leisure activity on the incidence of Alzheimer's disease. *Neurology*, 57(12), 2236–2242. <https://doi.org/10.1212/wnl.57.12.2236>
- [45] Scarmeas, N., Zarahn, E., Anderson, K. E., Habeck, C. G., Hilton, J., Flynn, J., Marder, K. S., Bell, K. L., Sackeim, H. A., Van Heertum, R. L., Moeller, J. R., & Stern, Y. (2003). Association of life activities with cerebral blood flow in Alzheimer disease: implications for the cognitive reserve hypothesis. *Archives of neurology*, 60(3), 359–365. <https://doi.org/10.1001/archneur.60.3.359>
- [46] Scarpina, F., & Tagini, S. (2017). The Stroop Color and Word Test. *Frontiers in psychology*, 8, 557. <https://doi.org/10.3389/fpsyg.2017.00557>
- [47] Schaub (2019). *Mental health books outsell diet and exercise books at Barnes & Noble*. Retrieved from <https://www.latimes.com/books/la-et-jc-mental-health-book-sales-20190111-story.html>
- [48] Schmand, B., Bakker, D., Saan, R., & Louman, J. (1991). The Dutch Reading Test for Adults: a measure of premorbid intelligence level. *Tijdschrift voor gerontologie en geriatrie*, 22 1, 15-9 .
- [49] Smith S. M. (2002). Fast robust automated brain extraction. *Human brain mapping*, 17(3), 143–155. <https://doi.org/10.1002/hbm.10062>
- [50] Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23 Suppl 1, S208–S219. <https://doi.org/10.1016/j.neuroimage.2004.07.051>
- [51] Staff, R. T., Murray, A. D., Deary, I. J., & Whalley, L. J. (2004). What provides cerebral reserve?. *Brain : a journal of neurology*, 127(Pt 5), 1191–1199. <https://doi.org/10.1093/brain/awh144>
- [52] Steffener, J., & Stern, Y. (2012). Exploring the neural basis of cognitive reserve in aging. *Biochimica et biophysica acta*, 1822(3), 467–473. <https://doi.org/10.1016/j.bbadis.2011.09.012>
- [53] Stern Y. (2006). Cognitive reserve and Alzheimer disease. *Alzheimer disease and associated disorders*, 20(2), 112–117. <https://doi.org/10.1097/01.wad.0000213815.20177.19>
- [54] Stern Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *The Lancet. Neurology*, 11(11), 1006–1012. [https://doi.org/10.1016/S1474-4422\(12\)70191-6](https://doi.org/10.1016/S1474-4422(12)70191-6)
- [55] Studenski, S., Carlson, M. C., Fillit, H., Greenough, W. T., Kramer, A., & Rebok, G. W. (2006). From bedside to bench: does mental and physical activity promote cognitive vitality in late life?. *Science of aging knowledge environment* : SAGE KE, 2006(10), pe21. <https://doi.org/10.1126/sageke.2006.10.pe21>
- [56] Sumowski, J. F., Wylie, G. R., Chiaravalloti, N., & DeLuca, J. (2010). Intellectual enrichment lessens the effect of brain atrophy on learning and memory in multiple sclerosis. *Neurology*, 74(24), 1942–1945. <https://doi.org/10.1212/WNL.0b013e3181e396be>
- [57] Tov (2019). *People spent \$1.9 billion last year on apps to keep their brains sharp as they age – here's what actually works*. Retrieved from <https://www.marketwatch.com/story/older-americans-spent-19-billion-last-year-on-apps-to-keep-their-brains-sharp-heres-what-actually-works-2019-05-24>
- [58] United States Census Bureau. (2019). *Older population and Aging*. Retrieved from <https://www.census.gov/topics/population/older-aging.html>
- [59] Walsh, K.W. (1991). *Understanding Brain Damage: A Primer of Neuropsychological Evaluation* (2nd. ed.). Edinburgh: Churchill Livingstone.
- [60] Wang, H. X., Jin, Y., Hendrie, H. C., Liang, C., Yang, L., Cheng, Y., Unverzagt, F. W., Ma, F., Hall, K. S., Murrell, J. R., Li, P., Bian, J., Pei, J. J., & Gao, S. (2013). Late life leisure activities and risk of cognitive decline. *The journals of gerontology. Series A, Biological sciences and medical sciences*, 68(2), 205–213. <https://doi.org/10.1093/gerona/gls153>
- [61] Wikee, G., & Martella, D. (2018). Capacidad física y reserva cognitiva como factores protectores de las funciones atencionales en adultos mayores [Physical activity and cognitive reserve as protective factors for attentional functioning in older people]. *Revista medica de Chile*, 146(5), 570–577. <https://doi.org/10.4067/s0034-98872018000500570>

AUTHORS

First Author – Michael Hess, M.Sc. candidate, Radboud University, mrhess25@gmail.com

Second Author – Joukje Oosterman, PhD, Radboud University, j.oosterman@donders.ru.nl