

UNIVERSITÀ
DI PAVIA

Dipartimento di Biologia e Biotecnologie
"L. Spallanzani"

Laurea Magistralis in Neurobiology

Downregulation of Bilateral Hippocampus and Default Network Subsystems during Retrieval Stopping: an fMRI Mega-Analysis

Supervisor:
Chiar.ma Prof.ssa Fulvia Palesi

Experimental thesis by
Federico Graziani

Academic Year 2025/2026

Thesis title

ENG Downregulation of Bilateral Hippocampus and Default Network Subsystems during Retrieval Stopping: an fMRI Mega-Analysis

ITA Deattivazione dell'ippocampo bilaterale e dei sottosistemi della Default Network durante il Retrieval Stopping: una mega-analisi fMRI

Abstract

Controlling our own internal world is an essential skill for adaptive, goal-directed behaviour. A domain-general inhibitory control mechanism, historically examined through stopping a prepotent motor response already underway, has also been found to block retrieval of memories that conflict with current goals. This ability underpins coherent engagement with the environment, is important for affect regulation and mental health, and is impaired in psychiatric disorders marked by intrusive thoughts, such as major depression, post-traumatic stress disorder, anxiety disorders, obsessive–compulsive disorder and schizophrenia.

While the suppression of internal thoughts can be challenging to investigate under ecological conditions, the Think/No-Think task provides an experimental paradigm to study participants' voluntary stopping of memory retrieval. Works on the neural correlates of this paradigm have established that prefrontal areas target the hippocampus, possibly resulting in neural inhibition within the hippocampal formation. Cortical modulations described in the literature further point to deactivation of the visual cortex, of stimulus-specific perceptual regions recruited during encoding of experimental items, and suggest a deactivation pattern in the Default Network.

However, no comprehensive quantitative account of the cortical correlates of Retrieval Stopping has ever been made.

In this project a mega-analysis pooling was conducted on raw functional Magnetic Resonance Imaging (fMRI) data acquired while humans perform the Think/No-Think task, to observe hippocampal and Default Network subsystems task modulation in a large sample. We found that both right hippocampus and left hippocampus were heavily responsive to task trials: activation during Think trials was comparable to baseline, whereas No-Think trials exhibited negative functional response, with a larger effect size in the right hippocampus relative to the left. Most interestingly, we found an activity dissociation of Default Network subsystems: while Default Network B (DN-B) activity is not modulated by the Think/No-Think task, Default Network A (DN-A) is strongly affected, exhibiting reduced activity during No-Think trials relative to Think trials. The results presented here show that both the right and left hippocampus are subject to deactivation during retrieval stopping, with the right hippocampus displaying a greater degree of modulation. They outline a pattern of cortical deactivation following retrieval stopping, highlighting how DN-A areas recapitulate hippocampal downregulation.

These findings bring further evidence for the functional specialization of Default Network subsystems: they corroborate current accounts of DN-A as underlying memory retrieval and internally directed thought, whereas DN-B's lack of modulation during the TNT task is consistent with its proposed role in non-mnemonic functions.

Abstract in Italiano

Il controllo del proprio mondo interiore è un'abilità essenziale per un comportamento flessibile, adatto per gli scopi del momento. E' stato dimostrato come un meccanismo di controllo inibitorio unico per qualsiasi dominio, storicamente studiato tramite l'interruzione di una risposta motoria già in atto, possa inibire anche il recupero di ricordi che entrano in conflitto con i nostri obiettivi attuali. Questa abilità è alla base di un'interazione coerente con l'ambiente, è importante per la regolazione affettiva e la salute mentale e risulta compromessa nei disturbi psichiatrici caratterizzati da pensieri intrusivi come la depressione maggiore, il disturbo da stress post-traumatico, i disturbi d'ansia, il disturbo ossessivo-compulsivo e la schizofrenia.

Nonostante la soppressione dei pensieri interni sia difficile da indagare in condizioni ecologiche, il paradigma Think/No-Think offre un modello sperimentale per studiare l'interruzione volontaria del recupero della memoria nei soggetti. Gli studi sui correlati neurali di questo paradigma hanno stabilito che le aree prefrontali bersagliano l'ippocampo, risultando potenzialmente a un'inibizione neurale della struttura. Le modulazioni corticali riportate in letteratura puntano inoltre a una deattivazione della corteccia visiva, delle aree percettive stimolo-specifiche attivate durante la codifica degli stimoli sperimentali, e suggeriscono un pattern di deattivazione della Default Network.

Tuttavia, non è ancora stata realizzata una valutazione quantitativa esaustiva dei correlati neurali nella corteccia del Retrieval Stopping.

Questa tesi riporta una mega-analisi condotta raggruppando dati di Risonanza Magnetica funzionale (fMRI) acquisiti durante l'esecuzione del compito Think/No-Think da parte di soggetti umani per analizzare la modulazione dell'ippocampo e dei sottosistemi della Default Network in un campione di grandi dimensioni. Abbiamo riscontrato che sia l'ippocampo destro che quello sinistro erano fortemente responsivi ai trial del compito: l'attivazione durante i trial Think era attorno alla baseline, mentre i trial No-Think mostravano una risposta funzionale negativa. Inoltre, l'effetto del trial era maggiore nell'ippocampo destro rispetto al sinistro.

Ancora più interessante, abbiamo riscontrato una dissociazione funzionale dei sottosistemi del Default Network: mentre l'attività del Default Network B (DN-B) non è modulata dal task Think/No-Think, il Default Network A (DN-A) è fortemente influenzato, mostrando una riduzione di attività nei trial No-Think rispetto ai trial Think.

I risultati qui presentati mostrano che sia l'ippocampo destro che quello sinistro sono soggetti a deattivazione durante il Retrieval Stopping, con modulazione più ampia nella struttura destra. Inoltre, delineano un pattern di deattivazione corticale evidenziando come

le aree del DN-A ricapitolino la deattivazione ippocampale. Queste scoperte portano quindi ulteriori evidenze a favore della specializzazione funzionale dei sottosistemi della Default Network: supportano la visione del DN-A alla base del recupero della memoria, mentre la mancata modulazione del DN-B durante il task è coerente con il suo ruolo proposto nelle funzioni non mnestiche.

Contents

Abstract **3**

Abstract in Italiano **5**

1 Introduction **9**
1.1 Retrieval Stopping 9
1.1.1 The Study Of Cognitive Mechanisms to Understand the Brain . . 9
1.1.2 Think/No-Think Task 10
1.1.3 Neural Correlates of the TNT Task 13
1.2 The Hippocampus 17
1.2.1 Historical Introduction 17
1.2.2 Hippocampal Functions 18
1.2.3 Hippocampal Inhibition In the Think/No-Think task 20
1.3 The Default Network 22
1.3.1 Historical Introduction of the Canonical Default Network 22
1.3.2 Fractionation 25
1.3.3 Functional Profile 26
1.4 Magnetic Resonance Imaging 27
1.4.1 functional Magnetic Resonance Imaging 28
1.4.2 fMRI diffusion and limitations 29

2 Aim Of The Work **30**

3 Materials and Methods **33**
3.1 Datasets and Participants 33
3.2 Study Design 36
3.3 Fmri Data Acquisition 37
3.4 Fmri Data Preprocessing 38
3.4.1 Realignment 38
3.4.2 Slice-Timing Correction 39
3.4.3 Coregistration and Segmentation 40
3.4.4 Normalization 40
3.5 fMRI Data analyses 41
3.5.1 First-Level Native Space Pipeline 41
3.5.2 Modelling 42

3.5.3 Region Of Interests 43
3.5.4 Signal To Noise Ratio 45
3.5.5 Percent Signal Change 45
3.6 Correlation Between Deactivations 47
3.7 Statistical Tests 47
3.8 Contrast Maps 48

4 Results **49**
4.1 Signal To Noise Ratio 49
4.2 Region Of Interest Analyses 49
4.2.1 PSC in Hippocampi 49
4.2.2 PSC in Default Network Subsystems 52
4.3 Correlation Between Deactivations 55
4.4 Qualitative Inspection Of Deactivations 56

5 Discussion **59**

6 Conclusions **67**

Bibliography **69**

Acknowledgements **81**

1 Introduction

1.1 Retrieval Stopping

Everybody has, at some point, felt the need to avoid remembering something once confronted with its reminder. We stumble back to places where something distressing happened, meet acquaintances who remind us of embarrassing events, or hear a song that reminds us of an unwelcomed memory. Remarkably, we seem to have the ability to voluntarily intervene when such a memory begins to surface into awareness: we are able to interrupt its retrieval even if it is strongly associated with the cue in front of us.
Intuition may suggest otherwise: we all know the experience of a thought becoming incessant into our minds the moment we are told not to think about it.
Nonetheless, we now have more than 20 years of research demonstrating not only that retrieval can be successfully blocked before the memory reaches awareness, but also that doing so has behavioural consequences on the memory trace, that it has considerable variability across the population, and that it relies on identifiable neural correlates.
The robustness of these findings has led to the characterization of a novel cognitive process, termed “Retrieval Stopping”.

1.1.1 The Study Of Cognitive Mechanisms to Understand the Brain

We might question the importance of studying cognitive processes at all: if neural activity can be directly recorded in animal models, and measured non-invasively in humans, why should we devote attention to theories of possible cognitive operations?

One answer comes from the history of human neuroscience: the field is derived from cognitivism, an information-processing approach to the study of the human mind that takes into account internal mental states and their operations to explain behaviour.
Cognitivism was formed in opposition to behavioralism, which, by restricting the discipline to observable stimuli and responses, was the major experimental psychology paradigm until the 1980s. In the history of the field advanced intuitions about the workings of the human mind have been obtained by abstract functional description, sustained by relatively simple behavioural experiments. For example, Donders, as far back as 1868, was able to infer the temporal sequencing of mental operations with measurement on reaction times while participants performed various experimental tasks [1].
Ebbinghaus, in 1885, pioneered the rigorous experimental study of memory retention and forgetting by quantifying learning curves using lists of nonsense syllables [2].

Tolman, in 1948, was able by studying rats to demonstrate that animals form internal ”cognitive maps” of their spatial environments, and are able to store them and use them to flexibly modify behaviour [3].

The emergence of neuroimaging techniques in the 1980s and 1990s had a massive effect on the field: the ability to scientifically observe human brain activity in real-time, non invasively, was such a step forward that a recently formed discipline quickly prevailed: cognitive neuroscience.
But the advent of functional neuroimaging does not result in an omniscient view of the workings of the human brain. Results on neural activations in mice or areas activated during certain tasks in humans only yield neural correlates of a phenomenon, not an explanation. A theoretical framework is needed if we want to advance knowledge on the human mind. Indeed, to understand operations performed by the neural tissue we must start to confront activity in the brain during different types of tasks conducted by the participants. To interpret these comparisons, we must understand what cognitive operations each task involves: we shall know how internal states, or symbols, are sustained, transformed and integrated, and only then we can postulate how these are reflected in computations by the neural tissue.
Data collection without a guiding theory risks to become mere “stamp collecting”: it is not rare to find accounts of neural correlates of a mental state even where it’s unclear what that specific state entails, or what the observed activity actually means. Cognitive theory provides a functional context for constraining measurements, manipulations and tasks, and to make final observations interpretable.

Finally, as much as theory can give us a framework, we shall not mistake a framework for reality: the metaphor of the brain as a computer can be powerful and have high explaining power, but it still remains a metaphor.

1.1.2 Think/No-Think Task

Over past decades, extensive research on action stopping, which is the process of how organisms stop physical actions already under way, has given rise to a voluminous literature on inhibitory control, often studied through the Go/No-Go task. Inhibitory control is postulated as a putative mechanism through which the brain suppresses representations or processes that conflict with an individual’s goals [4].
Initially, discussions about the mechanisms of thought stopping have built on the study of this action-stopping inhibitory control. However, the explanatory limits of the action stopping model regarding impairments in thought stopping, such as intrusive thinking in various psychiatric disorders, were evident. If workings of our internal world were to be

studied in a scientific manner, a novel paradigm was much needed.

The first account of executive control on declarative memories is the 2001 paper by Anderson and Green [5]. In this seminal work, the researchers introduced the Think/No-Think task (TNT task), a modification of the Go/No-go paradigm used to study executive control over motor actions. In this first version of the task participants were asked to learn word-pairs, so much that they would recall a member of each pair when provided with the other member.
Then, in the actual TNT phase, a cue from one of the pairs appeared on the computer screen. Depending on which cue appeared, subjects were told either to recall and say out loud the associated response word, or not to think about the response and not utter any sound. They first learned which cues would require suppression so that they would recognize these words on sight.
Subjects fixated on the cue word for the entire time that it appeared on the screen, generating a constant threat that the associated memory might intrude into consciousness.
Thus, suppression trials required the stopping of both a prepotent motor response and the entrance of an unwanted memory into awareness.
The objective of this first study was to determine whether attempting to prevent awareness of an unwanted memory would hinder its later retrieval: in a next phase, subjects had to recall the correct associate for each of the cue words. The researchers found a reduction in later recall of the suppressed items, relative to baseline items (items that were studied but neither suppressed nor retrieved), named Suppression-Induced Forgetting (SIF).
What's interesting is that SIF was present in recalling items even when they were probed with novel cues that had not been part of the TNT task, a feature named cue independence. Behavioural effects of reduced recall of No-Think items entails either a disruption of the mnemonic associations, or of the representations themselves. Cue independence speaks to the latter: because forgetting extends to cues that were never used during the TNT phase, Retrieval Stopping appears to act on the target memory itself rather than on a single retrieval route, although a small contribution of weakened cue–target associations cannot be ruled out [5].

On subsequent formulations of the Think/No-Think task items no longer have to be spoken out loud, and the condition assigned to each cue, either suppression or retrieval, is not learned before the TNT task, but is suggested during cue presentations. The cues are presented on a computer screen with a bounding box: if the associated item has to be retrieved, the bounding box is colored in green. If the associated item is supposed to be excluded from awareness, the bounding box is colored in red. If the cues are words, they can be presented without a bounding box but in colored fonts. If the associate should be retrieved the word is in green, while if it is to be blocked from awareness the cue word is

in red.

The study phase has also been changed: in recent years, the participants first are presented on a computer screen cues immediately followed by their associate, for a number of presentations. Then, participants are asked to select the learned associate in a forced-selection task, followed by feedback on the choice. Cue-target associations are presented to participants until a desired accuracy has been obtained in all pairs.
Furthermore, even after this practice-assisted learning, inquiries have been implemented to confirm that the associations are studied effectively: most experiments working with this task in recent years implements a test at the end of the study phase. Further analyses are conditioned on this test, meaning that if participants failed to immediately recall an associate once presented with the cue, that pair was removed from the TNT phase, or that pair data was removed from final analyses.

During the actual TNT phase, given that the association was extensively studied, seeing the cue by the participants will elicit automatic retrieval. During Think trials, on top of this automatic retrieval, if the participant complies with instructions, there will be an effort to retain the memory in awareness, and an effort to retrieve the associate if automatic retrieval failed. The cognitive process involved in these trials is termed “active retrieval”.

Instructions given to participants before and during the experiment are key for eliciting the cognitive processes desired.
Instructions that experimenters provide to participants for the Think/No-Think task were recently standardized to have consistent procedures across different researchers and different centers. The standardized instructions report that experimenters need to tell participants to “to look at, comprehend, and attend to the Hint on the screen until it disappears from the screen”.
For Think trials, participants are asked “to look at it, think about the associated Response as quickly as possible, and keep it in mind the entire time the Hint is on the screen ”.
For No-Think trials, “to prevent the associated Response from coming to mind at all, not even for a second, and not even after the Hint disappears from the screen”. Furthermore, if the response happens to come to mind, participants are told to “actively push it out of their mind”, and that they “should not replace the response with any other word, image, or idea” [6].
If these instructions are complied with by participants, No-Think trials will elicit automatic retrieval of the target once confronted with the reminder, against the participant goals, and there will be an effort to stop retrieval. A conflict between a representation triggered by an imperative stimulus and an individual’s goal is supposed to elicit inhibitory

control. This is the cognitive mechanism termed Retrieval Stopping.

1.1.3 Neural Correlates of the TNT Task

The first paper that explored neural correlates of the TNT task is Anderson *et al.* from 2004 [7].
In an event-related fMRI design, while participants performed the TNT task, they found a predominantly frontal set of areas more active during No-Think trials than Think trials: bilateral dorsolateral prefrontal cortex (DLPFC) and ventrolateral prefrontal cortex (VLPFC), anterior cingulate cortex (ACC), pre-supplementary motor area and dorsal premotor cortex (preSMA and PMDr respectively) and the intraparietal sulcus (IPS); increased activation in these areas predicted memory inhibition (SIF).
Furthermore, DLPFC activity was related to memory disruption even when targets were tested with an independent cue: this indicates that DLPFC is contributing to inhibition of memory traces.

What researchers immediately saw was how closely areas activated during No-Think trials were similar to prefrontal regions related to stopping prepotent motor responses [8].
Stopping an initiated action is achieved through inhibitory control and the same mechanism was supposed to be involved in controlling thought.
In subsequent years the prefrontal cortex came to be regarded as the seat of inhibitory control [9].
Conjunction analyses found that the same prefrontal areas were active both for inhibitory control over action and thought [10]. Furthermore, classifiers trained to distinguish inhibitory control over action or action execution could also distinguish between active retrieval and Retrieval Stopping, and vice-versa.
Strikingly, the same paper was also able to demonstrate a property of this inhibitory control system: the same right anterior dorsolateral prefrontal cortex (aDLPFC) and ventrolateral prefrontal cortex (VLPFC) regions could shift their effective connectivity so that they could switch from modulating the primary motor cortex during action stopping to modulating the hippocampus during Retrieval Stopping.
The domain-general inhibitory control mechanism can thus flexibly change the target of inhibitory control in a goal-directed fashion [9].

The original work from 2004 had indeed found a higher activation in both right and left hippocampus during Think trials relative to No-Think trials, pointing at a reduced activity once Retrieval Stopping was engaged.
Bilateral DLPFC predicted activation dynamics in the right hippocampus regarding items that were subsequently remembered or forgotten, but it did not predict left hippocampus

activity [7]. Those were the first findings indicating that the hippocampus and DLPFC could interact during attempts to suppress recollection of memories, resulting in forgetting. Hippocampus inhibition during Retrieval Stopping will be better described in section 1.2.3.

Not only subcortical structures, but cortical areas too were shown to have higher activity in Think rather than No-Think trials, already in the 2004 work [7]
While most of the subsequent works have focused on the hippocampus, a few studies have pointed to cortical deactivations as well.
In an fMRI study conducted in 2012, areas found less active during No-Think trials relative to Think trials were : occipital lobe , dorsal posterior cingulate cortex, retrosplenial cortex , Left insula/Heschl's gyrus, medial prefrontal cortex, subcallosal cortex and frontopolar cortex [11].
These recapitulate areas historically associated with Default Mode Network, as it will be better described in chapter 1.3.
Furthermore, recent studies have corroborated a theory which predicts which cortical areas can be downregulated during Retrieval Stopping:
If a cue is able to immediately activate the representations of the target, the hippocampus may rapidly reactivate neocortical regions [12]. This rapid cortical reinstatement, possibly experienced as a conscious awareness of the associate, against the participant's goal, may then upregulate inhibitory control in parallel to the hippocampus and the cortical region.
This means that inhibitory control not only targets areas subserving the retrieval process, but also cortical areas that code for the perception of the memory content.
If participants suppress visual memories of objects, fusiform cortex and lateral occipital cortex show reduced activity [13].
If participants suppress memories of scenes, parahippocampal cortex activity is reduced, subsequent visual memories are less detailed, and a classifier algorithm reports reduced scene information during and after the suppression attempt [14].
If participants suppress unpleasant scenes, parahippocampal place area and amygdala show reduced activity [15].
The nature of the content that is being reinstated thus predicts which cortical areas will be targeted by inhibitory control: this postulate has been named "reinstatement principle".

What's interesting is that no direct long-range projections have been found from the PFC to hippocampus, and the majority of cortical projections to subcortical structures have an excitatory nature [16].
Various theories of poly-synaptic inhibition pathways have been proposed [17]:

One area focused by these hypotheses is the dorsal-anterior cingulate cortex (dACC) : Retrieval Stopping engages the dACC and activations in this area can predict SIF. Furthermore dACC seems to be involved in supporting conflict detection [18], and has strong, diverse connections with PFC and strong links with MTL [19]. It is thus in a position to be able to receive top-down excitatory inputs from aDLPFC and propagate that influence to target areas. The entorhinal gating hypothesis suggests that the dACC suppresses cortical inputs to the hippocampus by activating parvalbumin-containing inhibitory neurons in medial rhinal and parahippocampal cortices [20]. While this theory could explain overall reduced activity in the hippocampus, it does not explain the active disruption of hippocampal traces themselves.

A second hypothesis, termed thalamo-hippocampal modulation hypothesis, posits that dACC forms strong bidirectional connections with thalamic nucleus reuniens (RE), which would suppress hippocampal activity. A recent preprint shows how in mice and rats nucleus reuniens projections target GABAergic inhibitory interneurons in CA1 hippocampal region [21].
It's important to note that these two hypotheses are not mutually exclusive, and both mechanisms may operate in tandem.

Another hypothesis proposes instead that Retrieval Stopping can suppress a region in the basal forebrain, called the medial septal nucleus (MSN), as evidenced by the fact that this structure shows reduced activity during No-Think trials [22].
This structure has a double effect on the hippocampus: it can drive the theta oscillations required for memory retrieval [23], and also exert an inhibitory hold on hippocampal interneurons. If functions of the MSN are impaired, it would result in tonic inhibition of hippocampal principal cells and lack of hippocampal theta [24].
This hypothesis may explain also why basal forebrain lesions that include the MSN induce confabulation, which has been attributed to failures to suppress irrelevant memories [25]. It still remains under investigation how the PFC can alter the activity of MSN.

A final hypothesis is concerned not with direct inhibition, but with rapid termination of driving input to the hippocampus.
The hypothesis generalizes a mechanism known to occur for action stopping to thought stopping: a connection from the PFC to the subthalamic nucleus (STN), that is able to suppress thalamo-cortical drive [26]. Certain thalamic nuclei sustain drive to regions providing input to or receiving output from the hippocampus, or even to the hippocampus itself. By modulating STN, we could thus have a regulation of awareness of thoughts and memories.

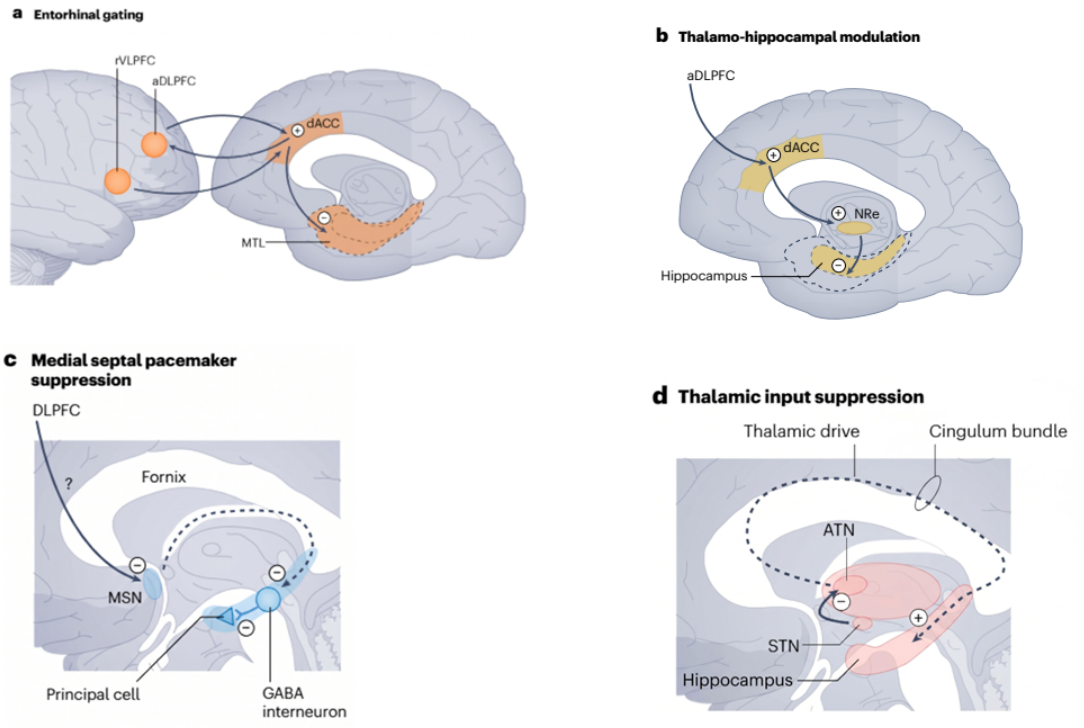


Figure 1: **Schemes of current hypothesis of hippocampal inhibition.** Depictions of brain sagittal slices with areas postulated to partake in the circuits highlighted.
a) Entorhinal gating hypothesis. rVLPFC = right ventrolateral prefrontal cortex.
aDLPFC = anterior dorsolateral prefrontal cortex.
dACC = dorsal anterior cingulate cortex.
MTL = medial temporal lobe.
b) thalamo-hippocampal modulation hypothesis. NRe = nucleus reuniens.
c) Medial septal pacemaker suppression hypothesis. MSN = medial septal nucleus.
d) Thalamic input suppression hypothesis. STN = subthalamic nucleus.
ATN = anterior nucleus of the thalamus.
Pictures adapted from Anderson *et al.* (2025) [27]

1.2 The Hippocampus

Much of the research on the Think/No-Think task has focused on the hippocampal structure. The hippocampus is an elongated structure within the medial temporal lobe that gets its name from historical gross dissections resembling a sea horse. It resides just below the 6-layer neocortex, and being an ancient brain structure, it is constituted of a three-layer archicortex [28].
Various subregions have been parcellated within the hippocampus proper: the cornus ammonis (CA), divided in three layers (CA1, CA2 and CA3), the dentate gyrus and the subiculum. While hippocampal neuronal circuitry is complex and still not completely understood, we can interpret the entorhinal cortex as the main input driver and the fornix as the main output hub [29].
Over the history of the study of this structure, we recognized that the hippocampus is vital for spatial navigation, learning, and the creation of new declarative memories [30].

1.2.1 Historical Introduction

After the early anatomical descriptions and dissections of the hippocampus in the 18th century, no clear function was attributed to this area. In the 1860s, the name of the structure became popular due to a scientific dispute between Thomas Henry Huxley, supporter of Charles Darwin's theory of evolution, and anatomist Richard Owen. Owen argued that the hippocampus minor was only present in human brains, with a supposed role of supporting human-unique mental properties. Huxley countered this with a successful dissection on an ape, demonstrating that the structure is indeed present in other animals [31].
Regarding the function of the structure, one theory that was developed in the 20th century and that persisted until the 1960s viewed the hippocampus as the seat of response inhibition: animals with hippocampal damage tended to be hyperactive, and had difficulty learning to inhibit previously learnt responses [32].
In the 1950s, a number of epilepsy patients or patients affected by schizophrenia proceeded with surgical lobotomy in an attempt to treat their debilitating symptoms.
In 1953 a patient later known as H.M underwent bilateral medial temporal lobe resection, having the anterior two-thirds of the hippocampus, entorhinal cortex and amygdala removed. What followed was an unexpected turn in the cognitive functions of the patient: H.M, could remember events that happened up to 3 years before the surgery, going back until his childhood, but could not remember events that happened the couple of years before the surgery. This was a case of partial retrograde amnesia, with a gradient through the lifetime of the patient.
Furthermore, the patient was unable to create new episodic memories subsequent to the

surgery, exhibiting a severe anterograde amnesia.
Studies of this peculiar cognitive functioning led to the hypothesis that a major function of the human hippocampus and its adjacent brain regions is to support the creation of new, declarative memories [33]. Over the next decades, the sparing of remote memories in the patient was further investigated: it pointed to a consolidation gradient, whereby memory traces could gradually become independent of the hippocampus, and transferred to another brain region.

1.2.2 Hippocampal Functions

What we know today regarding the peculiar amnestic profile that H.M. presented, is that the hippocampus does not store the content of long-term memories. According to the indexing theory, the hippocampus over time comes to storing a sparse "index" of the memory content, pointing to the neocortical patterns that had been active during the experience [34].
The hippocampus can rapidly form and retain an index representing the co-active neocortical zones thanks to Long-Term Potentiation, the proposed underlying biological mechanism. Subsequent reactivation of this hippocampal index by partial sensory cues reengages the original neocortical ensemble, reinstating the full conscious recollection of the episode. Subsequent evidence corroborated this theory, and the process came to be known as consolidation [35–37].

The most important computations that the hippocampal circuitry performs is thus not the ability to remember a single item, but the ability to bind together different sensory stimuli in a coherent context: the hippocampus is a system that organizes experience in relations, rather than in isolated traces [38].
It is clear how the relational framework provided by the hippocampus allows memories not only to be relived as experiences in the mind, but also underlies the ability of humans to construct a mental scene of possible scenarios, by puzzling together details from various memories to create a novel scene. The mental scenes can be either plausible futures or unrealistic scenarios, ability termed Mental Time Travel. The hippocampus thus seems to be involved in the creation of the unconstrained mental content, independent from current stimuli, that accompany us in every spare moment [39], and a study on encephalitis patients with hippocampal lesions found that while stimulus-independent thoughts would not stop with the presence of hippocampal damage, mental content would have a loss of vivid visual scenes about past, present or future, with a shift to more verbally mediated thoughts [40].

Although it may come across from the presented studies that the main role of the hippocampus is memory, a whole field of research has investigated another function subserved by the structure. An important function attributed to the hippocampal structure, other than memory, came from investigation of experimental lesions in laboratory animals. From early studies with experimental lesions in laboratory animals, it was reported how these specimens had an affected spatial navigation, especially in tests such as the Morrison Water Maze [41].
However, it was with technological advancements to detect real-time neural activity in-vivo in animals that the theory of the hippocampus as key for spatial navigation could go beyond the limits of lesion studies.
An important discovery after investigation of single cell recordings was the presence of pyramidal cells in cornus ammonis and granule cells in dentate gyrus which fired selectively when the animal was in a specific position within its environment: this discovery, other than securing a Nobel Prize to its researchers, suggested the theory that the hippocampus was able to create a cognitive map of the environment, and supported spatial navigation within it [42].
The discovery of these cells, aptly named "place cells", was followed by discovery of certain neurons in the medial entorhinal cortex which fired when the animal was in multiple precise locations in the environment, arranged in an hexagonal grid (grid cells) [43], discovery of border cells, firing when the animal is close to the boundary of the environment [44], and head-direction cells, which coded for the direction pointing by the animal's head [45].

Long debates have taken place trying to settle if the hippocampus is mainly a spatial mapping system or a learning/memory system: raw evidence could point to relations between memories and episodic details in respect to the subject being peculiar subtypes of navigation in a cognitive map, i.e. with mechanisms of memory having evolved from mechanisms of navigation in the physical world [46]. The opposite view instead conceives the only role of the hippocampus as coding relations between experiences in a coherent mental context, and postulates that spatial navigation is a peculiar application of relational memory organization [47].

While considerable research has tried to merge these two views in a unifying theory, a final solution remains elusive.

1.2.3 Hippocampal Inhibition In the Think/No-Think task

As quickly presented in chapter 1.1.3, the first neuroimaging study of the Think/No-Think (TNT) task showed a reduced activation of both hippocampi during No-Think trials relative to Think trials [7].
The TNT literature corroborated these results in numerous successive work, although across years a focus was concentrated on the right lateralized structure [11, 13, 15, 48–51]. Other than the contrast between conditions, studies found that activity during active retrieval would be higher than baseline (as a possible marker of mnemonic awareness), while, importantly, retrieval stopping would be reflected with reduced activity relative to baseline [11], termed Negative Bold Responses (NBRs). While these were quickly interpreted to be a proxy for inhibition, caution must be in place when interpretations jump from high-level functional neuroimaging results to underlying neural tissue dynamics. Some accusations have been made against the interpretations of NBRs: one accusation asked why, since inhibition in the cortex is often provided by GABAergic interneurons, Blood-Oxygenation-Level-Dependent signal would not increase, since these neurons would start to be metabolically active. Another accusation explains NBRs as methodological or technical artifacts of the functional imaging process, like a nearby area becoming more active and "stealing" blood from the area of interest [52].
Three main works have thus presented strong evidence to corroborate the hypothesis that the hippocampal structure suffers inhibition from the prefrontal control mechanism:

The first comes from an elegant study done in 2016 [53]. The authors modified the TNT task by presenting, between Think and No-Think trials, pictures of objects in scenes, called "bystander items". Bystanders presented both before and after No-Think trials were remembered much worse than those presented between Think trials, a deficit termed amnesic shadow. The fact that bystanders item suffered lower recall when they were presented both before No-Think trials (while they were being consolidated) and after No-Think trials (while they were being encoded) points to all functions of the hippocampus being disrupted, not only the retrieval of the target memory. This suggests that the whole structural area is affected, as in a transient virtual lesion, rather than a selective inhibition of the single suppressed trace. Furthermore, the magnitude of lower recall for bystanders items correlated across subjects with hippocampal NBRs, pointing to an inhibition of the whole structure.

Another one comes from the technique of magnetic spectroscopy, which is able to detect the concentration of specific neurotransmitters within a defined region in vivo. Schmitz *et al.* (2017) combined this technique with functional Magnetic Resonance Imaging during the TNT task. Resting hippocampal GABA concentration predicted the magnitude of

hippocampal NBR during No-Think trials and the amount of subsequent Suppression-Induced Forgetting.
Furthermore, effective connectivity showed that the right dorsolateral prefrontal cortex modulated the hippocampus only in participants with higher hippocampal GABA: the higher the inhibitory capacity of the hippocampus, the stronger its downregulation [22].

The last evidence does not come from Magnetic Resonance Imaging, but from EEG [54]. Theta power in the hippocampus supports hippocampal-cortical synchronization during retrieval, and we see reductions during No-Think trials, emerging around 500 ms after cue onset and lasting throughout the trial, localized to the medial temporal lobe and stronger in participants who forget successfully. EEG records directly from neural activity, independently of the hemodynamic response that makes NBRs ambiguous, proving how the hippocampal neural tissue is disrupted during Retrieval Stopping.

These findings, taken together, strongly corroborate the interpretation that hippocampal downregulation during Retrieval Stopping is inhibitory in nature.

1.3 The Default Network

1.3.1 Historical Introduction of the Canonical Default Network

Taking advantage of the different meanings of the umbrella term “network”, the Default Network relies its existence as a scientific object on the numerous methodological approaches used to discover it, without a proposal of a unified definition across functional, anatomical and psychological models [55]. Since Ingvar studies in the late seventies on regional Cerebral Blood Flow [56], various researchers were curious of understanding the importance of consistent, regionally-specific activity patterns found during rest states. With the adoption of Positron Emission Tomography (PET) scanning for brain imaging, higher resolution in observing the correlates of neural activity was achieved. The landmark paper by Shulman and colleagues [57] explored the riddle that had been found consistently across all early PET studies, i.e. a relative deactivation of certain rostral and caudal midline association cortical areas during the experimental task in respect to control resting-state recordings.
The seminal 2001 work from Raichle and colleagues found that this distinct set of regions were consistently deactivated across various disparate cognitive tasks: this led them to suppose the existence of coupled areas that would constitute a ”default mode of brain functions” [58]. Subsequent work was able to better determine this set of areas consistently deactivated: the early definition of the Default Network (also known as the Default Mode Network, DMN) was contrasting it with what it was considered not to be [59].
A parallel line of research was trying to understand if the areas found deactivated were in some ways connected to one another: In 2003, a paper by Greicius and Menon found that many of the brain regions that were consistently reported to be deactivated in PET studies showed synchronized, low-frequency fluctuations (high connectivity) [60]. The surge of these type of functional connectivity analysis on Magnetic Resonance Imaging data resulted in a rapid expansion of papers corroborating the existence of a coherent, intrinsic network of cortical regions defined named Default Mode Network, other than the presence of other “resting-state networks” [61,62]. A final methodology implemented to study, and define, the DMN was to observe functional activations. While most of the research up to this point had focused on task-related downregulation, a few studies were able to see task-evoked increases of activity above baseline in areas overlapping the network . Such tasks were self-referential judgments [63], autobiographical memory, episodic retrieval and social cognition [64].

We thus see how three different methodologies roughly recapitulate the same regions. These are oriented across a midline axis, with more caudal regions comprising poste-

rior cingulate cortex (PPC), retrosplenial cortex (Rsp) and ventral precuneus, and more frontal regions: ventral, anterior and dorsal medial prefrontal cortex (MPFC), and part of the basal forebrain. The parahippocampal cortex (PHC) in the middle temporal lobe region is also involved, as are lateral regions of the temporal lobe: lateral temporal cortex (LTC) and temporal pole (TempP).
In early accounts, also the hippocampal formation (HF+) was thought to be part of the network. Important regions reside in the posterior parietal cortex: both the temporo-parietal junction (TPJ) and the posterior-inferior parietal lobule (pIPL) are key for the network.
Furthermore, recent research is moving beyond a cortico-centric approach, appreciating the influence and participation in the network of subcortical areas such as amygdala, caudate, anterior and mediodorsal thalamic nuclei, and regions of the cerebellum [65].

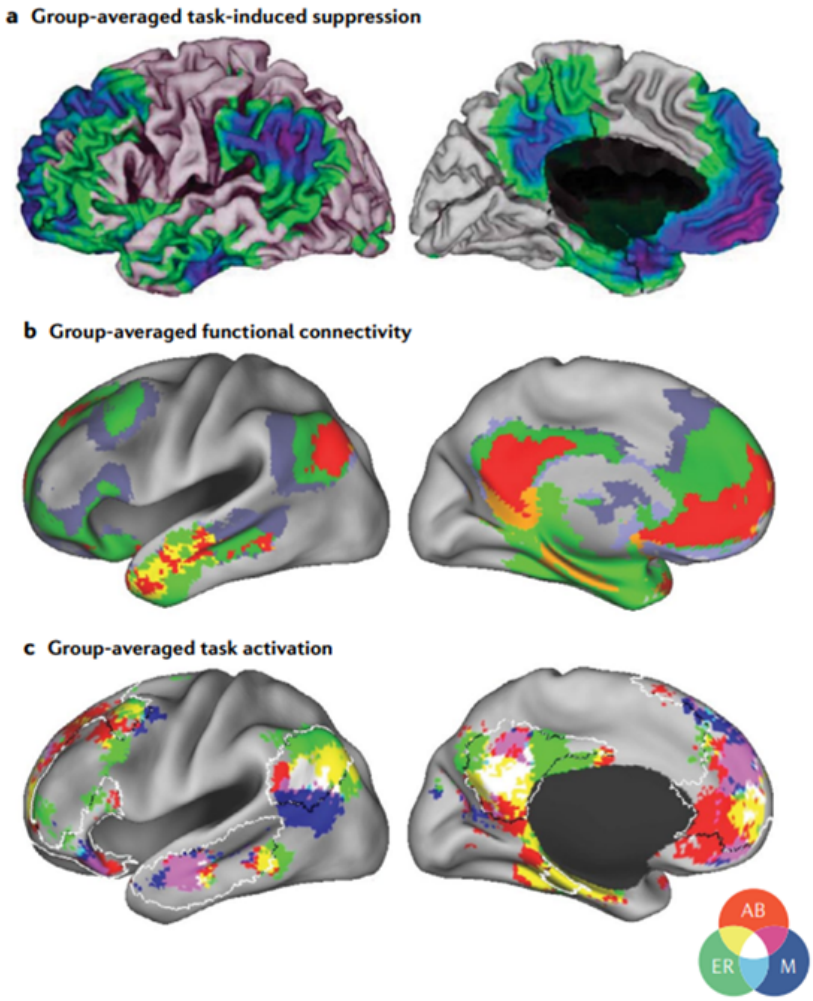


Figure 2: **Depictions of the three methodologies that have defined the canonical Default Network.**
a) Inflated brain with lateral and medial views of task-induced suppressions, highlighting caudal and rostral midline regions, tempo-lateral regions and posterior parietal cortex regions. Data from the original Raichle *et al.* 2001 study [58]. Purple indicates maximum deactivation.
b) Group-averaged functional connectivity data, showing overlap over functional connectivity maps with 3 different seeds: dorsomedial prefrontal cortex, ventromedial prefrontal cortex, hippocampus. Warm colors indicate maximum overlap between the three maps. Areas highlighted recapitulate task-induced downregulation areas.
c) Task-evoked activation and their overlap over three different cognitive tasks: autobiographical remembering (AB), episodic retrieval (ER) and social cognition requiring Theory Of Mind (M). High overlap areas overlap areas outlined in white, which are network boundaries derived from functional connectivity analyses.
Figures from Buckner & DiNicola (2019) [66].

1.3.2 Fractionation

In 2010, Andrews-Hanna and collaborators published a ground-breaking study that challenged the concept of a single monolithic Default Network. The paper performed hierarchical clustering on functional connectivity data of regions within the canonical network, and was able to support an account of a fractionation of the network: the results showed two distinct subsystems that interacted with a core set of hubs.
Furthermore, the two subsystems exhibited task-based functional dissociation , with core regions sharing functional properties of both subsystems.
The subsystem named "MTL" was composed of HF+, PHC, Rsp, ventral MPFC, and pIPL , and showed activation when future prospective decisions about one's self in the future were contrasted with decision in respect to present self.
The other subsystem, named "dMPFC" was instead composed of dorsal MPFC, TPJ, LTC, and TempP. This system was preferentially active while participants reflected on questions on themselves, in respect to questions on non-personal semantic knowledge. Interestingly, regions within the dMPFC subsystem are also activated when participants infer the mental states of other people, so a complex role at the interplay between mental state inference, social cognition, and the use of one's own mental states as a model for inferring the mental states of others was hypothesised.
PPC and anterior MPFC composed the shared core, as demonstrated by high betweenness centrality in graph measures and shared activation profile in tasks that would differentiate the two subsystems [67].
The functional specialization between these two subsystems was corroborated by numerous following accounts, linking the MTL subsystem to episodic retrieval and simulation, and the dMPFC subsystem to mentalizing [68–70].

A paradigm shift slowly happened from 2017. Braga and Buckner, (part of the same lab that published the 2010 work) noticed that all data regarding the exploration of the DMN relied on group-averaged data, which could be subject to artifacts and loss of spatial resolution.
They thus extensively scanned individuals (24 MRI sessions each) to obtain enough data to calculate functional connectivity with high precision within individuals. The authors were able to demonstrate that when within-subject data was inspected, the canonical Default Network fractioned into two distinct parallel networks that were closely interdigitated throughout the topography of many cortical regions. They named these two dissociable systems Default Network A and Default Network B. Default Network A was approximating areas of the MTL subsystem, while Default Network B lacked coupling with the parahippocampal cortex, similar to the dMPFC subsystem, although juxtaposi-

tion of the networks made the comparison nuanced.
The hypothesis was thus that while group-averaged data was showing a dissociation between two networks with the presence of a shared core, within-subject precision imaging would point at only two dissociable interdigitated systems. The closely interdigitated areas, when averaged across participants, would have been subject to spatial blurring of their signal, resulting in the illusion of a third independent activation profile, engaged across domain-specific tasks, that was interpreted as a core region shared between subsystems [71].

Other accounts have affirmed this fractionation, thanks to extensive within-individual precision imaging [72–74]. Furthermore, Du *et al.* 2024 created a parcellation atlas following this type of methodology, from which the present work extracted the topography of the Default Network subsystems for analyses. We thus followed their practice of using the name Default Network instead of Default Mode Network.

1.3.3 Functional Profile

Two different lines of research can be described that explore the effects of the canonical Default Network (DN) on cognition.
One is about the passive influence the DN has in cognition. Numerous compelling evidence shows how externally-focused attention inhibits the network, and how ineffective suppression of DN results in impaired selective attention, slower responses, and higher error [75].
The other focuses on active and direct roles of the DN on cognition, and can be grouped in the exploration of 5 functions: self-reference judgements, social cognition, memory retrieval and episodic construction, language processing and mind wandering [76].
This last function has particularly pervaded the history of the network. The search for the neural basis of the "streams of consciousness" as defined by William James in 1890 [77], of that stimulus-independent thoughts that accompany us during much of our awake time [78], has driven part of the DN research [79].

Regarding recent accounts of fractionation of the network in Default Network A and Default Network B thanks to within-individual precision imaging, results point to the two networks supporting different cognitive functions.
Evidence shows how the two can be dissociated by episodic projection task contrast (thinking about past self or future self in contrast with present self) and theory-of-mind task contrasts (thinking about another's false belief in contrast with assessing a false photo of the person), which preferentially recruit DN-A and DN-B, respectively.
DN-A therefore appears more related to scene construction and episodic simulation, whereas DN-B appears more related to mentalizing and social-cognitive inference [72, 80].

1.4 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI), originally called Nuclear Magnetic Resonance Imaging, is an imaging technique developed in the early 1970s motivated by cancer detection, exploiting the phenomenon of Nuclear Magnetic Resonance. MRI was superior to Computer Tomography scans due to better contrast in soft tissue images, and the absence of ionizing rays. It has become widely used for its non-invasive nature and good spatial resolution.
While Nuclear Magnetic Resonance can be applied to various nuclei with non-zero spin, MRI scanners are usually tuned to image hydrogen atoms.
In water, free hydrogen nuclei are characterised by an intrinsic magnetic momentum, arising from their spin, which is randomly oriented. With the presence of a strong magnetic field, in MRI scanners called B_0, hydrogen magnetic moment aligns in orientation with the direction of the magnetic field, precessing at a precise Larmor frequency.
A radio frequency pulse (RF) is then applied tuned to this Larmor frequency, with an electromagnetic radiation whose magnetic component B_1 is oriented at 90° in respect to B_0. The RF applies the B_1 magnetic field which alters magnetization vectors of hydrogen nuclei, making it oscillate on the plane on which B_1 lies.
When the RF pulse is concluded, the hydrogen nuclei are no longer driven by B_1, and the net magnetization returns to its equilibrium state. This process is called relaxation, and it follows two independent time courses.
The longitudinal magnetization, aligned with B_0, progressively recovers its initial value with a time constant T1, called spin-lattice relaxation. At the same time, the component of the magnetization vector along the B_1 plane decays with a time constant T2, called spin-spin relaxation, as the magnetic moments of neighbouring nuclei interact with each other and lose their phase coherence. The precessing transverse magnetization behaves as a rotating magnetic moment, and by Faraday induction it generates an oscillating current in the receiver coil: this creates the raw signal.

The raw signal does not yet carry spatial information. To obtain it, we create gradients in the static magnetic field for each direction, as Larmor frequencies of nuclei are proportional to field strength. We thus have signals only from nuclei that resonate with a known RF value, encoding their position across each gradient.

Based on the timing at which RF is given and terminated, and the timing of the recording before the process starts again, different tissues will return different signal levels. We can alter the timings of the scanner pulses and recording (sequences), in particular by changing the Repetition Time (TR) and Echo Time (TE), to weight differently the T1

time constant or T2 time constant, obtaining different levels of signals from tissues, being able to create various tissue contrasts. T1-weighted scanning is the standard technique for obtaining structural images of the soft tissue of the brain. Different sequences have been optimized for different functions and to image different biological processes [81].

1.4.1 functional Magnetic Resonance Imaging

Functional Magnetic Resonance Imaging (fMRI) is the technique of implementing a Blood-Oxygenation-Level-Dependent (BOLD) signal sensitive MRI sequence. Spin-spin relaxation, determined by the time constant T2, is sensitive to local fluctuations of the magnetic field experienced by the detected nucleus.
In practice, the signal decays faster than T2 would predict, with a time constant named T2*. What Seiji Ogawa discovered in 1990 is that this decay could be used to infer neural activity. Oxygen around the body is transported by hemoglobin, which once bound to an oxygen molecule is diamagnetic (oxyhemoglobin), and becomes paramagnetic once the oxygen molecule is lost (deoxyhemoglobin).
The paramagnetic properties of deoxyhemoglobin accelerate T2* decay: as more deoxyhemoglobin is near a certain hydrogen nucleus, more signal is lost at the echo time in a T2*-sensitive sequence due to local field inhomogeneities.
Neurons in the brain are among the most metabolically active cells in the body, and thus need plenty of oxygen to survive. Through a complex interaction with astrocytes, more oxygen and nutrients are given to more metabolically active neurons in a localized manner, with a delay of just a few seconds, via an increase in blood volume and perfusion, a process known as neurovascular coupling. The dynamic of the BOLD signal following localized neural activity is described by the Hemodynamic Response Function (HRF): early oxygen consumption by the more metabolically active neurons creates an initial dip in the BOLD signal. The vascular response, with a delay of 3 to 4 seconds, surges the area with oxygenated blood, more than is actually needed, so deoxygenated hemoglobin decreases in concentration, increasing the signal. The signal peaks at around 6 seconds after initial surge in activity, at which the deoxygenated/oxygenated hemoglobin ratio is the lowest. As cerebral blood flow is reduced, the BOLD signal starts to slowly fall, as the hemoglobin ratio returns to baseline levels.
With a proper MRI sequence it is possible to observe increases and decreases of the BOLD signal, and to model the response function to have a proxy of underlying neural activity [82].

1.4.2 fMRI diffusion and limitations

fMRI has been widely used since its conception. Both in its clinical use and research use it has transformed the scientific fields in which it has been applied [83]. If during the year 1995 127 papers were published in medical journals on fMRI results, 30 years later, in 2025, the papers published are 6363 (PubMed query).
It is superior to the other functional neuroimaging technique widely spread during the birth of fMRI: Positron Emission Tomography (PET). In respect to PET, it does not require a tracer injection in the patient's body, so can be regarded as a non-invasive technique, and does not require a nearby cyclotron laboratory to create radioactive tracers. It has superior spatial resolution, and can have images with good contrast between tissues.
Limitations of the technique are various:
First of all, the loud noises, small gantry, mandatory instructions to be still and restriction on metallic implants, make the technique suited for only a subset of the population. Low temporal resolution due to slow hemodynamic response makes the technique ill-suited to explore fast cognitive and neural mechanisms, as thousands of neural events and computations can be made in the timeframe used.
Spatial resolution used to be a major limitation, but we know that with higher static magnetic field strengths, the net magnetization can provide a stronger signal, achieving smaller voxels. Cognitive research thus shifted from using standard clinical scanners at 1.5 Tesla, which provided isotropic voxels of around 3mm, to having the strongest MRI scanners in the world at 11.7 Tesla, which can have resolution down to 1.2mm [84].

Another main problem of fMRI is careful interpretation of results. First of all, we have to remind ourselves that we are not recording direct neural activity: we are recording changes in blood oxygenation, as a proxy for neural activity. Furthermore, the phenomenon that links these two, i.e. neurovascular coupling, is still not completely understood. The other setback on interpretation is the high abstraction from raw signal to final results. What most pipelines observe is not BOLD signal directly, but a scaling weight for BOLD signal related to a certain condition in a linear model.
A careful interpretation of results is needed also to not fall into the error of performing a reverse inference. While thanks to study after study high accuracy can be achieved in predicting the neural correlates of a psychological state, the opposite is rarely true. To calculate the likelihood of the participants experiencing a certain mental content given the neural activation, Bayesian calculus must be performed. We should not reach quick conclusions, given that most brain regions are engaged by many different processes.

2 Aim Of The Work

The present study was conducted together with the Memory Research Group at the Medical Research Council Cognition and Brain sciences Unit (MRC-CBU) in Cambridge, UK. The objective was to elaborate a comprehensive characterization of hippocampal neural activity and cortical neural activity as registered with BOLD fMRI during Retrieval Stopping. While numerous accounts reported small downregulations in the hippocampus related to Retrieval Stopping, effect sizes varied across different papers.
Furthermore, a gap in the literature is present: while some work points to cortical deactivations in Retrieval Stopping relative to active retrieval recapitulating areas of Default Network, no proper exploration has ever been realized.
We thus conducted a mega-analysis on functional Magnetic Resonance Imaging datasets acquired in past years in three different neuroimaging centres. Raw functional images were used as primary data, and we were able to pool 330 participants' data.
We assessed functional activity in the brain by calculating Percent Signal Change in predetermined parcellated areas. We extracted data from the hippocampus thanks to structural masks manually traced for each individual, with the left and right hippocampus traced separately. We focused on activity of the Default Network as defined by recent fractionation of this network thanks to precision imaging. We thus implemented two different masks for the two subsystems that seem to compose this network: Default Network A and Default Network B.

We expect to observe reduced activity in both hippocampi during No-Think trials relative to Think trials, and to observe Negative Bold Responses during No-Think trials.
Even if some past work had shown hippocampal positive activity during active retrieval [11], we do not expect PSC to be higher during Think trials relative to baseline in a big sample.
No past work directly compares right and left hippocampus effect sizes in a large sample, but we suppose the right hippocampus to exhibit a bigger effect.

We expect Default Network A to be modulated by the task, and to recapitulate hippocampal behaviour. We thus expect it to have a lower activity during No-Think trials relative to Think trials, as we believe that inhibitory control might target this network in parallel with the hippocampus.
We expect Default Network B to show a small-to-none modulation, as it should not participate in the memory retrieval process, so it should not be targeted by inhibitory control, but functional coherence between two subsystems might still be present.

We expect the two networks to behave as functionally coherent across regions within them, and to exhibit functional specialization.
Furthermore, we expect to have correlation between magnitude of the Retrieval Stopping effect on hippocampal activity and magnitude of the effect in Default Network A activity across subjects, as both systems are hypothesised to be targeted by the same control mechanism.

3 Materials and Methods

3.1 Datasets and Participants

A multi-centered mega-analysis was conducted on functional Magnetic Resonance Imaging (fMRI) raw time-series as primary data.
Ten datasets were included in the present analysis, for a total of 330 human subjects. Datasets were not publicly available, and were accessed thanks to agreements between researchers who conducted the experiments and Professor Michael Anderson, manager of the current project. Datasets were coded with two digits cardinal numbers from 01 to 10, and were all gathered from 2012 onwards. Datasets 01,02 and 04 were collected at Lewis Centre for Neuroimaging, University of Oregon, Eugene, Oregon, U.S.A. Datasets 03,05,06,07,08 were collected at MRC Cognition and Brain Science Unit, Cambridge, UK. Datasets 09 and 10 were collected at Southwest University, Beibei, Chongqing, China.
All participants had no reported history of neurological, medical, visual, or memory disorders, and had normal or corrected-to-normal vision.
Datasets ranged in participants scanned, from 16 up to 137 participants. Participants sex was intended to be counterbalanced, although some datasets presented disproportions. All participants in all datasets were young adults.
Additional demographic information for each dataset can be found in table 1. Limited demographic information was available for dataset 01 and 02.

Dataset code	N (M/F)	Mean age (age range)	Scanner & location	EPI sequence	Items	Conditions modelled	Unmodelled data points (percentage)
01	18 (8/10)	(18 – 33)	Lewis Centre for Neuroimaging 3T Siemens Allegra	TR = 2 s, TE = 25 ms 34 interleaved axial slices 3.125 x 3.125 x 4 mm voxel size flip angle 80°	Words pairs	T / NI / I / Err / Rate	42.11%
02	18	N/D	Lewis Centre for Neuroimaging 3T Siemens Allegra	N/D	Word pairs	T / NI / I / Err / Rate	38.53%
03	16 (8/8)	25.6 ± 4.87	MRC CBU 3-T Siemens TIM Trio	TR = 2 s, TE = 30 ms 32 descending axial slices 3 × 3 x 3 mm voxel size flip angle = 78°	Word–picture pairings.	T / NI / I	52.52%
04	24 (8/16)	22.4 ± 3.76	Lewis Centre for Neuroimaging 3T Siemens Allegra	TR = 2 s, TE = 30 ms 31 interleaved slices 3 x 3 x 3 mm voxel size	Face-scene pairs	T / NI / I / nuisance	44.20%
05	27 (11/16)	22.9 ± 3.44	MRC CBU 3T Siemens Magnetom Prisma	TR = 1.12 s, TE = 30 ms 60 interleaved axial slices 2 x 2 x 2 mm voxel size	Word pairs	T / NI / I / Err / Rate / NULL	36.79%
06	18 (6/12)	23.7 ± 2.28	MRC CBU 3T Siemens TIM Trio	TR= 2 s,TE= 30 ms; 32 descending axial slices 3 x 3 x 3 mm voxel size flip angle 78°	Word pairs	T / No-Think/ Err	44.39%

Dataset code	N (M/F)	Mean age (age range)	Scanner & location	EPI sequence	Items	Conditions modelled	Unmodelled data points (percentage)
07	24 (12/12)	22.7 ± 0.50	MRC CBU 3T Siemens TIM Trio	TR = 2 s, TE = 30 ms 32 descending axial slices 3 × 3 × 3 mm voxel size	Word-object pairs	T / No-Think / Err	53.99%
08	24 (7/17)	24.2 ± 4.50	MRC CBU 3T Siemens TIM Trio	TR = 2 s, TE = 30 ms 32 descending axial slices 3 × 3 × 3 mm voxel size	Word pairs	T / No-Think / Rest / nuisance	58.88%
09	24 (12/12)	21.6 ± 1.98	Southwest University 3 T Siemens Trio	TR = 1.5 s, TE = 29 ms 25 interleaved axial slices 3 x 3 x 5 voxel size flip angle = 90°	Words pairs (Chinese characters)	T / No-Think / nuisance	55.86%
10	137 (33/104)	20.1 ± 1.83	Southwest University 3 T Siemens Trio	TR = 2 s; TE = 30 ms 32 interleaved slices 3.4 × 3.4 x 3 mm voxel size flip angle = 90°	Word pairs (Chinese characters)	T / No-Think / Fix / Err	4.48%

Table 1: **Demographics available for each dataset, scanning specs, items and condition modelled.**
TR = Repetition Time. TE = Echo Time.
Conditions: "T" = Think trials. "NI" = No-Think trials with no reported intrusion. "I" = No-Think trials with intrusion reported. "Err" = trials without participants' response during response time, and all trials with items that were not learned during study phase. "Rate" = time spent by participant rating if an intrusion occurred or not. "Fix" = time spent by participants looking at a fixation cross, "nuisance" or "NULL" = other variables of no interest.
Unmodelled data points for each dataset represent data points in which all GLM regressor had a zero value, divided by all data-points for all scanning sessions, and transformed in percentage. There was a consensus in this measure between participants in each dataset.

3.2 Study Design

All datasets implemented slightly modified versions of the Think/No-Think (TNT) task, originally devised by Anderson and Green in 2001 [5].

Items were diverse across the 10 datasets. Seven datasets used written words as both cues and targets. Other items used were word-faces associates, word-items, face-scenes, and word-objects. Nine datasets out of ten used emotional neutral stimuli, while dataset 4 used both neutral and emotional negative pictures. To assess the variability of neural activation elicited by different items, we employed statistical tests with datasets as a between-subject factor in all of our analysis. Supplementary information about items used can be found in table 1.

All the items were assigned to the suppress (No-Think), recall (Think), and baseline conditions, counterbalanced across participants. As described in 1.4.1, participants first went through an initial learning phase: this phase was designed to create strong associations between two items, so much that the target would likely intrude into awareness involuntarily once confronted with its reminder. Participants then engaged in the actual TNT phase in the MRI scanner, while fMRI BOLD signal was being recorded.
On each trial, participants received a cue on a screen in front of them: trials in Think condition were characterised by cues in a green bounding box, or were written in green. Participants were asked to pay full attention to the cue and retrieve the associated memory as quickly as possible, keeping it in mind while the reminder remained onscreen.
By contrast, in the No-Think condition, characterised by cues in a red bounding box or written in red, participants were asked to pay full attention to the cue and avoid looking away. They were asked to prevent the associate word from entering awareness during the whole trial, without engaging in any distracting mental activity. This "direct-suppression" instruction emphasized that the associate was not to be replaced with alternative words or thoughts. If the target came to mind anyway, they were asked to "push it out of mind".
Each cue was presented consistently in either Think condition or No-think condition.

All datasets gathered data with a final test phase, to explore behavioural effects on memory of Retrieval Stopping. We were not interested in behavioural results in the present study.

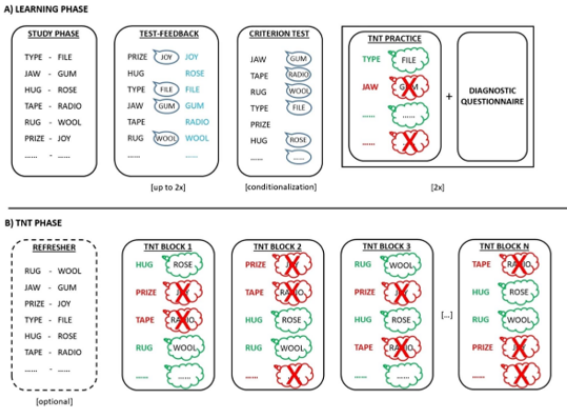


Figure 3: **Schematics of the learning phase and TNT phase of the Think/No-Think task.** After participants studied associations in the first study phase, the test-feedback phase was prolonged until a desired accuracy was obtained on all items. Analyses were performed only on items for which the criterion test was passed.
Here cue and target are both words, but other stimuli can be used.
An optional refresher was presented in certain datasets before the actual TNT phase. In the picture here words are presented as either written in green or red, while certain studies presented a bounding box around them. Figure from Nardo & Anderson (2024) [6].

3.3 Fmri Data Acquisition

All data analysed in this work were gathered through Magnetic Resonance Imaging at a field strength of 3 Tesla. All scanners were produced by Siemens (Siemens Healthineers,Erlangen, Germany), and different model information used by each dataset can be found in table 1.
All functional images were gathered during TNT phase with blood oxygenation level-dependent (BOLD) contrast, through a T2*-weighted gradient-echo echo-planar sequence (EPI images), with sequence parameters varying in each dataset.
Repetition Time (TR) varied between 1.2 seconds to 2 seconds. Echo Time (TE) varied between 25 milliseconds to 30 milliseconds. Voxels resolutions varied between 2 mm^3 isometric voxels to 3 x 3 x 4 mm in dimension. To assess the variability of results given different scanning parameters, we again employed statistical tests with datasets as a between-subject factor in all our analysis. Sessions (runs) per subject varied between datasets, ranging from 4 to 6.

Structural scans were collected with magnetisation-prepared, rapid gradient echo (MPRAGE) T1-weighted images. Voxel resolution was $1mm^3$ isometric voxels for all datasets except dataset 4, in which structural scan voxel resolution was 0.5 x 0.5 x 1 mm in dimension.
Additional information on scanning parameters can be found in table 1.
It is worth noting the significant differences between the 10 datasets, as scanner type,

acquisition sequence, voxel size, population background and task instruction can all alter final results. These differences allowed us to test the robustness of our conclusions in this mega-analysis.

3.4 Fmri Data Preprocessing

Preprocessing, first level and second level analyses of EPI images were conducted through MATLAB scripts (The MathWorks, Inc., Natick, MA, USA), calling relevant functions from Statistical Parametric Mapping software v12 release (SPM12; Functional Imaging Laboratory, UCL Queen Square Institute of Neurology, London).
Raw functional data were used as primary data for data pooling, to conduct a mega-analysis. Data pooling in this sense can maximise statistical power, and by allowing aggregation of sub-significant effects across independent studies, can effectively reduce both false positive and false negative results [85]. Furthermore, we could thus assemble a more representative participant sample of the general population.
Preprocessing steps conducted were realignment, slice-timing correction, coregistration, segmentation and normalization. First level analyses were conducted in native space, and Percent Signal Change was calculated for the main analyses. Second level group analysis in Montreal Neurological Institute (MNI) space was conducted to corroborate main results.

Each preprocessing step was conducted for each scanning session (run) independently, separately for each subject in each dataset. The pipeline described here was applied with equal parameters to every dataset, to obtain a standardized set of data to feed to first-level analyses.

3.4.1 Realignment

To address the misalignment between images in a time-series that occur due to head motion, a realignment step was applied (also known as motion-correction). Bulk motion can have large effects on final results, as it can create artifacts on BOLD activation maps [86]. Realignment assumes that head motion can be described using a rigid body transformation, following translation and rotation along each of the three axes. The session-specific mean EPI image was chosen as reference for realignment: SPM12 first realigned every session image to the first image of the session acquired, then calculated the mean volume from all the session volumes. The program then realigned every raw image to this calculated mean image. This was done by first temporarily smoothing images with Full Width at Half Maximum (FWHM) 5mm Gaussian kernel, then sampling voxel values ev-

ery 4mm. Each value sampled was confronted with values sampled from the reference image, and a least-squared cost function was applied to find the optimal transformation parameters. Images were then respliced to a new voxel grid using 4 dimensional B-spline functions. Head-motion translation and rotation parameters were saved to conditionalize following analyses, as excessive motion in one session could create artifacts not compensated by standard realignment tools.

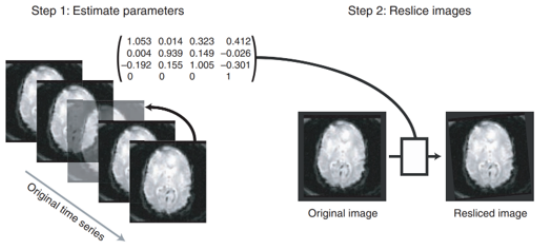


Figure 4: **Overview of the realignment process.**
The first step entails calculating the optimal transformation parameters to lower the cost function, to find the best transformation of each volume to the target (in this picture the middle one). The second step then creates a resliced version of each image following the parameters obtained in the first step. Figure from Poldrack *et al.* (2011) [86].

3.4.2 Slice-Timing Correction

fMRI works with 2D slices acquisition, meaning that data in different parts of the volume are acquired at systematically different times within a Repetition Time (TR). These differences in the acquisition time of different voxels, ranging up to several seconds, can create a mismatch with statistical models that expect a task-evoked signal in a precise time-point, and assume simultaneous data collection. Even if past evidence advised that slice-timing correction could be optional, especially if interleaved acquisition order was adopted [86], we decided to apply this step in our pipeline. The choice was mainly dictated by the fact that most datasets analysed in this work implemented a descending slice acquisition order, and by recent evidence showing how this step is critical in an fMRI pipeline [87]. Slice-timing correction requests a target slice as reference: for every subject for every dataset, the middle slice in the acquisition order was chosen. SPM12 then temporally interpolates the data of all the other slices with a sinc function to extract a predicted voxel value at the time the reference slice was taken.

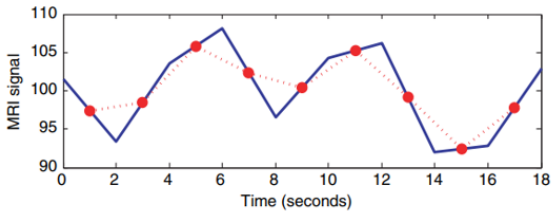


Figure 5: **A simple illustration for understanding slice-timing correction.**
A voxel time course with data points every 2 seconds (TR = 2) is respliced to values taken at TR/2. The depiction shows linear interpolation for simplicity, while normally a sinc function is used. Figure from Poldrack *et al.* (2011) [86].

3.4.3 Coregistration and Segmentation

The T1-weighted image was coregistered to the mean EPI, meaning that rigid-body transformation parameters were estimated to align the structural image to the average functional volume. A Mutual Information cost function was used to handle the between-modalities transformation. Once the T1 image was coregistered, every subsequent transformation or warping of the T1 image could easily be applied to EPI volumes too.
The structural image was corrected for bias field inhomogeneities, then segmented in individual tissues: grey matter, white matter, Cerebro Spinal Fluid (CSF), skull tissue, soft tissue outside the brain, background.
It was then normalized with SPM default non-linear warping to MNI space as implemented in SPM12 (MNI152Nlin6Asym template).

3.4.4 Normalization

DARTEL procedure was implemented for normalization, computing a flow-field to obtain parameters to transform individual subject coordinates to MNI space. Importantly for this project, both forward (y) and inverse (-y) parameters were saved for following steps.

3.5 fMRI Data analyses

3.5.1 First-Level Native Space Pipeline

The main analyses were performed in native space, meaning that analyses on time courses from subjects' voxel grids were applied directly after preprocessing on subject-specific coordinates, in each subject independently. No normalization was thus conducted to a common atlas, and no smoothing was applied to EPI images.
We were interested in BOLD signal modulation by the TNT task in the hippocampi, conducting a mega-analysis on the effect of Retrieval Stopping on these structures. We were also interested in Default Network A and B, as defined by recent evidence on the fractionation of the Default Mode Network.

The choice for a native space pipeline was twofold:
The first aim was to obtain the maximum precision about the anatomical position of our signal. We were interested specifically in Default Network subsystems and hippocampal formations, and group-average data can affect spatial precision of the signal. Each individual brain is not equal to everyone else's, but has peculiar structures and idiosyncrasies, accounting for anatomical variability. To compute results in a sample of participants in a group-based study, each individual brain is warped to a common atlas, employing computationally complex non-linear transforming functions.
After this process, there's still a risk of not-complete alignment between individual activation loci and atlas position, which could affect anatomical position of final results.
Furthermore, even after perfect structural alignment to a common atlas, each subject can show activation of functional areas in loci slightly different from the rest of the population, a phenomenon known as functional variability, which can again affect anatomical precision of final results. The standard approach to avoid these inaccuracies is to implement smoothing, that is the convolution of the three-dimensional image with a three-dimensional Gaussian filter (called a kernel).
The second reason a native-space pipeline was adopted was exactly to avoid the smoothing procedure. Voxels belonging to DN A subsystem are closely juxtaposed and surrounded by voxels belonging to DN B subsystem, described as an interdigitation phenomenon. By smoothing EPI images, we could risk BOLD signal from one area leaking into another area, affecting final results.

3.5.2 Modelling

We applied a General Linear Model (GLM) to subject-specific preprocessed BOLD signal time series. Functional sessions with excessive head motion were excluded from the analysis. The threshold was set at any single data point in the session exhibiting a translation parameter greater than 3mm or a rotation parameter exceeding 3°.
We submitted the model onset times of key experimental conditions and their duration; they were then converted to a box-car function, which is a discontinuous function that has a zero value except for specific intervals where it is equal to a constant value. Then, all box-car functions were convolved with the canonical hemodynamic response function (HRF), modelled in SPM with a double gamma function, to approximate the idealized BOLD triggered by underlying neural activity.

Datasets varied in how key conditions were modelled: datasets 1 to 5 modelled separately No-Think trials with intrusion reports and No-Think trials with no intrusion reported. All datasets modelled separately Think trials.
fMRI analyses were conditionalized from the criterion phase, so Think and No-Think onset times corresponding to items that participants failed to learn during study phase were excluded from the main regressors and grouped in the regressor of no interest. Datasets varied in nuisance regressors implemented.
Additional information on conditions modelled can be found in table 1. Each model included within-session global scaling (SPM default scaling), high-pass filtering using a cutoff frequency set at 1/128 Hz to remove low-frequency trends, and removal of temporal auto-correlation through pre-whitening procedure with AR(1) model. This last step is particularly important as standard optimization methods for GLM assume independence between one data point and another. Since the BOLD signal has a slow timecourse, we must remove the temporal autocorrelation between datapoints before applying optimization algorithms.

Six further regressors contained the head-movement parameters obtained during spatial realignment and were included as covariates (three translation parameters, one for each axis, and three rotation parameters, one for each rotating direction).

If, for example, we investigate a dataset which has 5 functional runs per subject, and models 6 experimental conditions, the final design matrix for one subject would comprise 65 columns, assuming no sessions are excluded. This architecture arises because each session is modeled independently, yielding 12 run-specific regressors per session (six convolved task conditions and six motion covariates) alongside five session-specific constant terms placed at the end of the matrix to account for the mean baseline signal of each individual run, termed intercepts.

Once the design matrix was constructed, the GLM scaled each regressor to best approximate the observed BOLD signal time course, and saved the optimal scaling parameter for each key regressor, called a β weight. This estimation was performed for each voxel independently using a mass-univariate approach. Because temporal autocorrelation was regressed out via the AR(1) process, the parameters were estimated employing Weighted Least Squares, rather than standard ordinary least squares. This algorithm calculates the optimal β weights by minimizing the sum of the squared residual errors between the model's predicted time course and the actual, pre-whitened fMRI data [88].
We thus obtained a distinct β weight map for each experimental condition, providing a quantitative measure of the task-driven BOLD response amplitude at every voxel, independently for each subject.

3.5.3 Region Of Interests

Region Of Interest (ROI) based analyses were conducted to estimate effect size of task modulation avoiding exploratory-analysis bias [89]. Hippocampal ROIs were manually traced on subjects' T1 structural images for all 330 subjects. This was done as a joint effort of lab members, following Pruessner *et al.*'s protocol [90]. Areas explicitly included were Cornus Ammonis (CA) regions, dentate gyrus, subiculum, alveus, fimbria, and part of the fascicular gyrus adjacent to CA. The right and left hippocampus were defined separately. A manual visual inspection was conducted on a randomly sampled subset of participants to assess integrity and spatial localization of hippocampal ROIs.

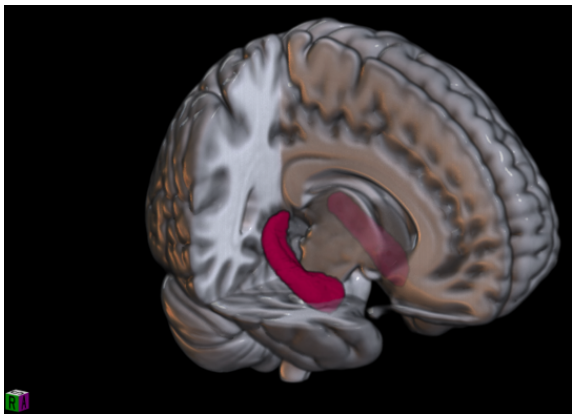


Figure 6: **3D render of the MNI structural template with an overlaid probabilistic mask of bilateral hippocampus,**
provided only for visualization purposes. The mask was generated by warping each individually traced hippocampal mask to MNI standard space and computing the spatial overlap across all subjects. The resulting probability map was then converted into percentage and thresholded at a 50% spatial concordance. Render created with MRIcroGL software [91].

To define Default Network ROIs we took advantage of recent breakthroughs in cortical networks parcellation through extensive imaging. We obtained masks from the Du15NET Consensus map, an atlas created through precision individual neuroimaging on 15 subjects. Every participant was extensively scanned in resting state for up to 11 sessions on different days, a functional connectivity matrix was calculated for 81,924 vertexes, and these were clustered to obtain 15 resting-state functional networks. The authors applied a novel multisession hierarchical Bayesian model (MS-HBM) algorithm to estimate the within-individual cortical networks: this algorithm was implemented as its critical feature was the ability to separate day-to-day within-subject noise to inter-subject differences [92].
The final consensus map represents the assignment of each vertex to the most common network observed across the 15 individuals collected [74].

Binary masks were thus obtained of Default Network A and Default Network B subsystems in MNI 152 space. To obtain masks in subjects' native space, we warped the binary masks through the DARTEL function following the inverse of the subject-specific flow-field derived from normalizing the T1 structural image (-y parameters), with Nearest Neighbour interpolation applied to preserve binary values.

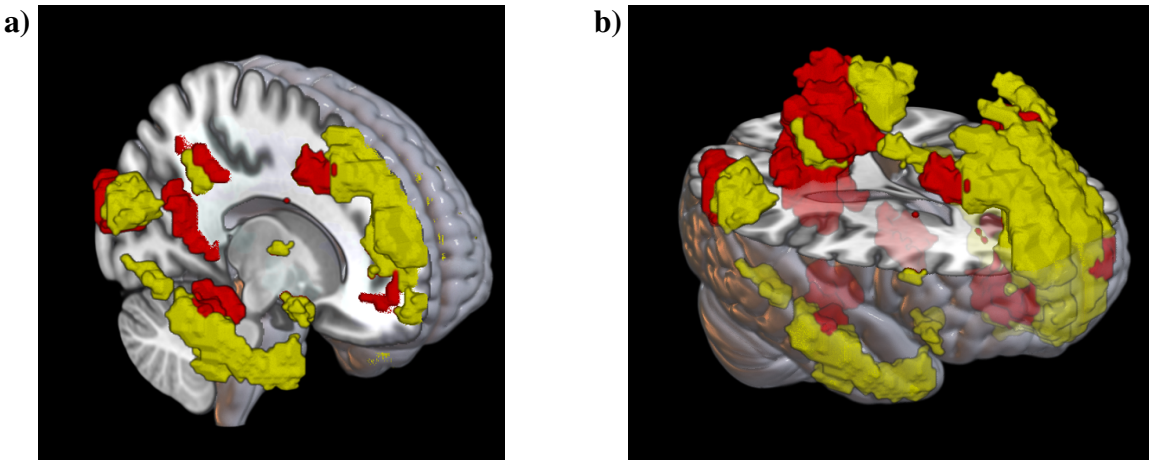


Figure 7: **3D renders of the MNI structural template with overlaid binary masks of Default Network Regions Of Interest (ROIs).**
Default Network A is presented in red, and Default Network B is presented in yellow. Masks are presented here in MNI space before warping procedure to each subject's brain for visualization purposes.
a) A sagittal cut has been made to observe lateral and medial regions, with masks obscured by structural voxels.
b) A horizontal cut has been made to better appreciate dorsal regions, with structural brain transparent to present all ROI voxels simultaneously. We can observe how in medial regions, both caudally and rostrally, various voxels of one mask are surrounded by voxels of the other mask, with closely juxtaposed areas. Renders created with MRIcroGL software [91].

3.5.4 Signal To Noise Ratio

As said before, one of the reasons researchers apply a smoothing step of EPI images during preprocessing is to account for spatial functional variability after group-averaging all subject brains, but this did not pertain to us as analyses were conducted in native space.
Another reason smoothing is recommended is due to the fact that the BOLD signal is inherently noisy, and signal extracted from single voxels can be overwhelmed by random noise: smoothing can enhance Signal to Noise Ratio (SNR) [93]. Since our preprocessing did not entail this step, it was crucial to check that SNR was sufficient in our data before continuing with the rest of the analyses. For fMRI BOLD signal data the standard quality metric is temporal Signal to Noise Ratio (tSNR). In early years of the field tSNR was calculated by dividing the temporal mean of a voxel's time series by its standard deviation across the whole time course. Whilst valid, this approach artificially increases the metric by treating task-driven BOLD fluctuations as noise. We opted for calculating residual tSNR, which is calculated by dividing the temporal mean signal of a voxel by the standard deviation of the residuals after GLM fitting [94].
Let a voxel time course be described as a total of N data-points, and after GLM fitting, let e_t represent the unexplained residual noise. The standard deviation of the residuals, σ_e, is calculated as:

$$\sigma_e = \sqrt{\frac{1}{N-p}\sum_{t=1}^{N} e_t^2}$$

where p represents the degrees of freedom (total number of regressors) estimated by the GLM.
The residual temporal Signal-to-Noise Ratio for each voxel is therefore defined as:

$$tSNR_{res} = \frac{\bar{Y}}{\sigma_e}$$

Once tSNR maps were obtained, they were thresholded at a value equal to 40, regarded in the literature as a threshold for quality signal [94]. Then they were projected as masks on subject-specific structural scans, to observe areas which had an insufficient tSNR, in MRIcroGL software [91].

3.5.5 Percent Signal Change

Raw β weights from first level analysis are subject to high between-subject variability. While having a big sample size should allow us to have sufficient statistical power, it would be optimal to work with less variability of our dependent variable.

To work with a more standard measure, we computed Percent Signal Change (PSC).
PSC of a defined region, be it a single voxel or a cluster of voxels, can be simply understood as the magnitude of the BOLD response divided by the overall mean of the BOLD time series. For our ROIs, we decided to apply a mean PSC approach. This means that all the time courses of voxels within a region of interest were averaged together to create a mean time course for that ROI. Another GLM was then applied to the mean timecourse, with the same design matrix as the first-level GLM. The procedure of averaging thousands of voxels in a single time course could create complications in the interpretation of results, which will be assessed later.

Percent Signal Change for a given condition is normally calculated as PSC = $100 \cdot \beta \cdot$ BM/MN, where β is the value of the parameter estimate from the GLM for that condition (beta weight) , BM is the distance from the baseline-to-maximum height of the regressor, and MN is the mean of the BOLD time series [86]. Problems arise with this formulation, as BM, the maximum height of the regressor in the time series, is dependent on subject-specific variables, experimental design specifications and stimuli used. Comparing PSC across different studies would become impossible, rendering it a meaningless metric of communicating the magnitude of a BOLD response. Following Poldrack's advice, "one should use the baseline to max distance of a specified event type in the calculation and report the specific event type and canonical HRF used when reporting percent signal change" [86]. For this reason, in our PSC estimation we calculated BM as the height of a standard double gamma HRF (as implemented in SPM12) following a theoretical instantaneous stimulus. BM thus becomes the standardized baseline-to-maximum height of a specific trial type, serving as a ruler for the calculation.
We calculated mean PSC in each ROI for every condition of interest, for each functional session independently.

For datasets 1 to 5 we averaged PSC for Intrusions and Non-Intrusion, to obtain a corresponding value of the No-Think condition. We averaged PSC for every session, thus obtained PSC for each condition for each of the 4 ROIs, in every subject.

We then explored Negative Bold Responses, challenging if the mean PSC across subjects was different from baseline. Importantly for this measure, in our first level analysis not all data points in the functional time courses were modelled by regressors in the GLM. The baseline for each voxel is thus constituted of the average BOLD signal level during all unmodelled data points. Prevalence of unmodelled data-points is described in table 1. It follows that in our mean PSC approach, the baseline for a given ROI is the average activity of all included voxels during all unmodelled data points.

3.6 Correlation Between Deactivations

Difference of PSC between Think and No-Think condition was computed for each ROI for all subjects. Given the observed functional coherence between the right and left hippocampus, we averaged the difference between the two ROIs in each subject.
We then calculated Pearson's correlation coefficient between the difference of activity between the two conditions in hippocampi (ΔPSC for both hippocampi), and difference of activity in Default Network A, across subjects (ΔPSC DN-A). We then did the same between the difference of activity in hippocampi and difference of activity in Default Network B (ΔPSC DN-B), across subjects. Scatter-plots with best-fit lines were constructed for visualization.

3.7 Statistical Tests

A 2x2 repeated measures ANOVA with ROI and trial type as independent variables and PSC as dependent variable was conducted to explore trial type effect, ROI effect and their interaction. ROI variable levels were hippocampus left and hippocampus right, and trial-type levels were Think and No-Think conditions.
Another equal 2x2 repeated measures ANOVA with ROI and trial type as independent variables and PSC as dependent variable was conducted to explore effects in Default Network Subsystems. ROI variable levels were Default Network A and Default Network B, and trial-type levels were Think and No-Think conditions.
Dataset type was added as a between subject factor in both ANOVAs, to account for heterogeneity of variance.
ANOVAs and t-tests both assume that residuals of data in respect to condition-specific average follow a Gaussian distribution. We implemented Quantile-Quantile (QQ) plots to check for normal distribution. No outliers were removed from our data.
Post-hoc tests were conducted to explore trial-type effects in each ROI separately, implementing two-tailed paired t-tests. To assess Negative Bold Responses, that is BOLD activity below baseline, two-tailed one-sample t-tests were performed for each condition in each ROI separately. Pearson's correlation coefficient between differences in activity significance was calculated with Student t-test.
All analyses were performed on JASP software (JASP Team, (Version 0.96.0) 2026).

3.8 Contrast Maps

A second pipeline was constructed for a qualitative corroboration of results and for visualization purposes. It consisted in warping the preprocessed EPI images to MNI space following subject-specific DARTEL flow-fields derived from the normalization step of the T1 structural image.

No smoothing kernels were applied to EPI images for whole-brain univariate analysis. After modelling with the aforementioned GLM, contrast maps were obtained for each subject. A second level analysis was conducted with an ANOVA to obtain the group-level difference between Think and No-Think conditions ($T > NT$). Dataset-specific covariates were included in the model to account for scanning parameters heterogeneity, while inter-subject variance was accounted for by the within-subject structure of the ANOVA. Cohen's d maps thresholded at a small effect size (Cohen's d = 0.2) were obtained, and projected in MNI template in MRIcroGL software [91]. The p-value map was calculated from the t-statistic parametric map of $T > NT$ contrast, applying a Family-Wise Error (FWE) corrected threshold of $p < 0.01$, combined with a cluster forming threshold of k=10.

4 Results

4.1 Signal To Noise Ratio

Residual temporal Signal to Noise Ratio (tSNR) was calculated for every voxel in each subject, volumes were thresholded at quality threshold recommendation of 40, and plotted in MRIcroGL software [91] on subject specific brain. Randomly sampled participants were manually checked in MRIcroGL to check if voxel within our Regions of Interest (ROIs) had insufficient tSNR.
After extensive sampling, no subject was found to have tSNR under a value of 40 in our ROIs.

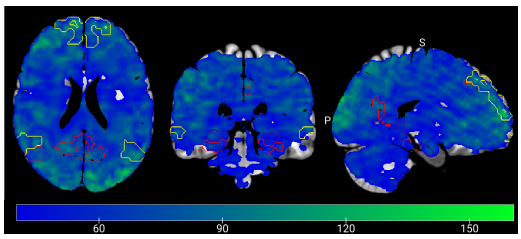


Figure 8: **Typical presentation of temporal Signal to Noise Ratio map**, thresholded at a value of 40 and thresholded with background, overlaid on T1 weighted structural scan.
Colored lines trace edges of Regions Of Interest (ROIs), red for Default Network A, yellow for Default Network B, and purple for hippocampal formations.
Slices taken at $X = 18.6$, $Y = -33$, $Z = 21$ through MRIcroGL software [91].

4.2 Region Of Interest Analyses

To compute a standard measure of BOLD activity in the selected ROIs, we calculated mean Percent Signal Change (PSC) associated with the Think condition and No-Think condition.

4.2.1 PSC in Hippocampi

Q-Q plots generated from the aggregated residuals of the 2x2 ANOVA model for Hippocampal Formations PSC were evaluated to assess the assumption of normality. A deviation from normality was present in the lower distribution tail, while the vast majority of the data closely followed a normal distribution. Given the large sample size and the established robustness of the ANOVA to minor normality violations, we decided to proceed with ANOVA tests and subsequent t-tests [95].

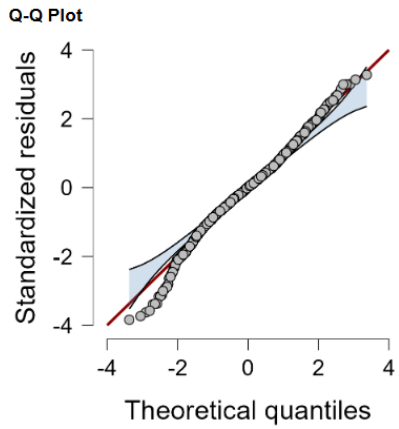


Figure 9: **Q-Q plot of residuals of Hippocampal Formation Percent Signal Changes (PSC).** Residuals are plotted against a theoretical normal distribution (diagonal red line). Light blue bands represent the 95% confidence intervals. Data points are 330 subject PSC for each of the 4 conditions, for a total of 1320 data points.

In analyses conducted in hippocampal ROIs, we found a significant trial-type $*$ ROI interaction, with a small effect size ($\eta_p^2 = 0.022$, $p = 0.008$).
We conducted post-hoc tests to test for trial-type effects in each ROI and observe the direction of the interaction.
We found that No-Think PSC was reliably lower than Think PSC for the left hippocampus, with a small effect size (Cohen's $d = 0.273$, $p < 0.001$). The same was found for the right hippocampus (Cohen's $d = 0.343$, $p < 0.001$).
The interaction found is thus due to effect size for the right hippocampus being larger than the left hippocampus.

We then explored Negative Bold Responses (NBRs).
The PSC of the Think condition was not statistically different from zero in both ROIs, with $p = 0.388$ and $p = 0.345$ for hippocampus left and right respectively.
Statistically significant differences were found instead for No-Think PSC, it being reliably lower than baseline both for the left hippocampus and the right hippocampus (Cohen's $d = -0.310$, $p < 0.001$, Cohen's $d = -0.365$, $p < 0.001$ for left and right respectively).

These results confirm past literature, showing how Think trials do not affect hippocampal activity in respect to baseline, while No-Think trials result in reduced activity in respect to baseline.

We did not find a significant dataset effect on trial-type$*$ROI interaction ($p = 0.45$, $\eta_p^2 = 0.027$)

Raincloud plots were obtained from the same data to better appreciate the activation distribution of group averaged data. Given the large sample size, activation distribution closely resembles a Gaussian curve. We can see how green distribution curves, plotting Think trials data points, are centered around zero (baseline), while red distribution curves, plotting No-Think trials,are centered on more negative PSC values.

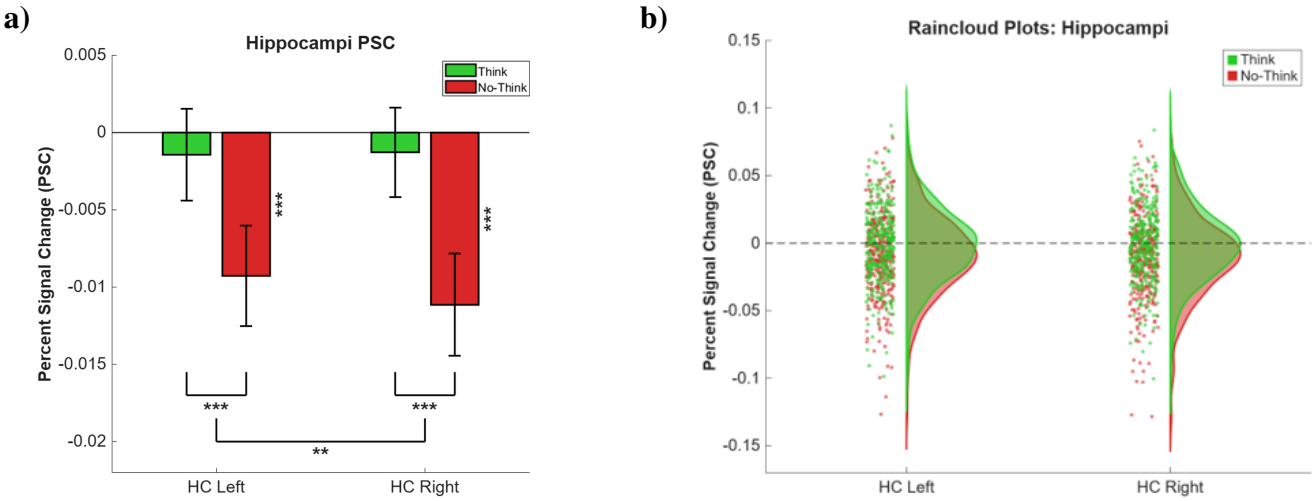


Figure 10: **Hippocampal Percent Signal Change (PSC) results.**
a) Bar graphs for hippocampal PSC. On Y axis PSC is presented, while on X axis Region Of Interest and Trial Type variables are shown. Green bars represent PSC associated with Think condition, while red bars represent PSC associated with No-Think condition. Total ANOVA sample size $N = 330$. HC Left = left hippocampus, HC Right = right hippocampus.
Error bars represent condition-specific 95% confidence interval.
Statistical significance is shown as: $p < 0.01$ (**), $p < 0.001$ (***).
b) Raincloud plots from the same data, with datapoints for each individual on the left-hand side and density distribution on the right-hand side, for Percent Signal Change associated with Think and No-Think conditions in the left hippocampus (HC Left) and the right hippocampus (HC Right).

4.2.2 PSC in Default Network Subsystems

Q-Q plots generated from the aggregated residuals of the 2x2 ANOVA model were again evaluated to assess the assumption of normality for Default Network subsystems ROIs. Again, small deviations from normality were present, here in both lower and upper distribution tails. As before, we proceeded with the ANOVA and subsequent t-tests.

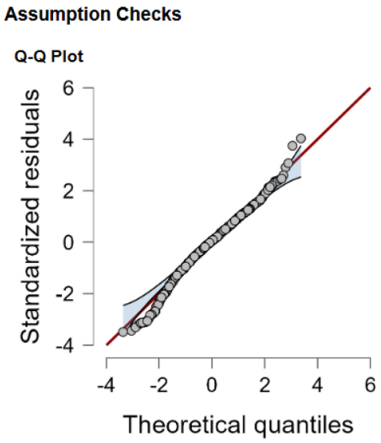


Figure 11: **Q-Q plot of residuals of Default Network Subsystems Percent Signal Changes,** plotted against a theoretical normal distribution (diagonal red line).
Light blue bands represent the 95% confidence intervals..

Analyses for Default Network subsystems showed a large trial-type∗ROI interaction ($\eta_p^2 = 0.563$, $p < 0.001$).
We again conducted post-hoc tests: we found that No-Think PSC was statistically significantly lower from Think PSC in Default Network A, with a large effect size (Cohen's $d = 0.761$, $p < 0.001$).
No statistically significant difference was found for Default Network B ($p = 0.615$).
The large interaction effect was thus a result of Default Network A showing modulation for trial-type, while Default Network B didn't exhibit any trial-type effect.

We then explored the difference of PSC from baseline, exploring Negative Bold Responses.
We found that all four conditions showed Negative Bold Response ($p < 0.001$ in each).

Raincloud plots were again modelled to appreciate distribution of activations.
Interstistingly, we see how PSC distribution for the Think condition (green) in DN-A does not resemble a perfect bell curve. It seems better described as a bell curve with a shoulder centered around a PSC value of - 0.05, that deviates from normality. This is interesting, as the rest of the distribution seems centered around zero. This could mean that the Negative Bold Response found in previous analyses for Think condition in DN-A could be driven

by only a subset of participants, while the activity of the rest of the sample is centered around a baseline activation, as we see in hippocampi distribution.

We can then appreciate how the Think and No-Think trials distributions differ in DN-A, with the No-Think distribution (red) centered on more negative values, while the two distributions are coincident for DN-B.

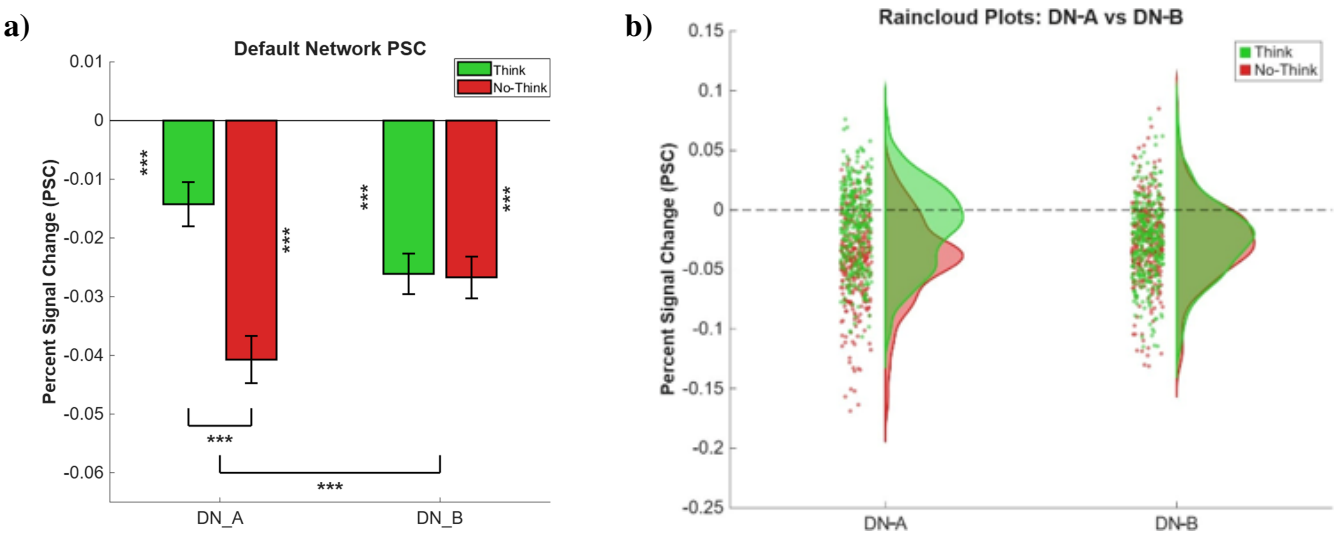


Figure 12: **Default Network subsystems Percent Signal Change (PSC) results.**
a) Bar graphs for hippocampal PSC. On Y axis PSC is presented, while on X axis Region Of Interest and Trial Type variables are shown. Green bars represent PSC associated with Think condition, while red bars represent PSC associated with No-Think condition.
Total ANOVA sample size $N = 330$.
Error bars represent condition-specific 95% confidence interval.
Statistical significance is shown as: $p < 0.001$ (***)
b) Raincloud plots from the same data.
DN-A = Default Network A DN-B = Default Network B

We did find a significant dataset effect on trial-type$*$ROI interaction, with a small effect size ($p < 0.028$, $\eta_p^2 = 0.056$).
We thus calculated PSC for each dataset independently.
Following t-tests performed in each ROI for each dataset separately, eight dataset out of ten reiterated Default Network A effect found in the main analysis, as No-Think PSC was statistically significantly lower than Think PSC.
Most datasets also recapitulated Default Network B behaviour, as it was modulated in 3 out of 10 datasets, two of which showed a crossover interaction (dataset 03 and dataset 07), that is, No-Think PSC was statistically significantly higher than Think PSC. This corroborates main findings, showing that No-Think trials do not result in decreased activity in DN B in respect to Think trials, yet result in decreased activity in DN A.

Furthermore, in 2 datasets out of 10 (dataset 08 and dataset 10) Think condition PSC was not statistically different from baseline, similar to what was found in hippocampi. This was observed from within-condition 95% Confidence Intervals. It is intriguing to observe that this effect was particularly present in dataset 10, which had a massive sample size (137 participants).

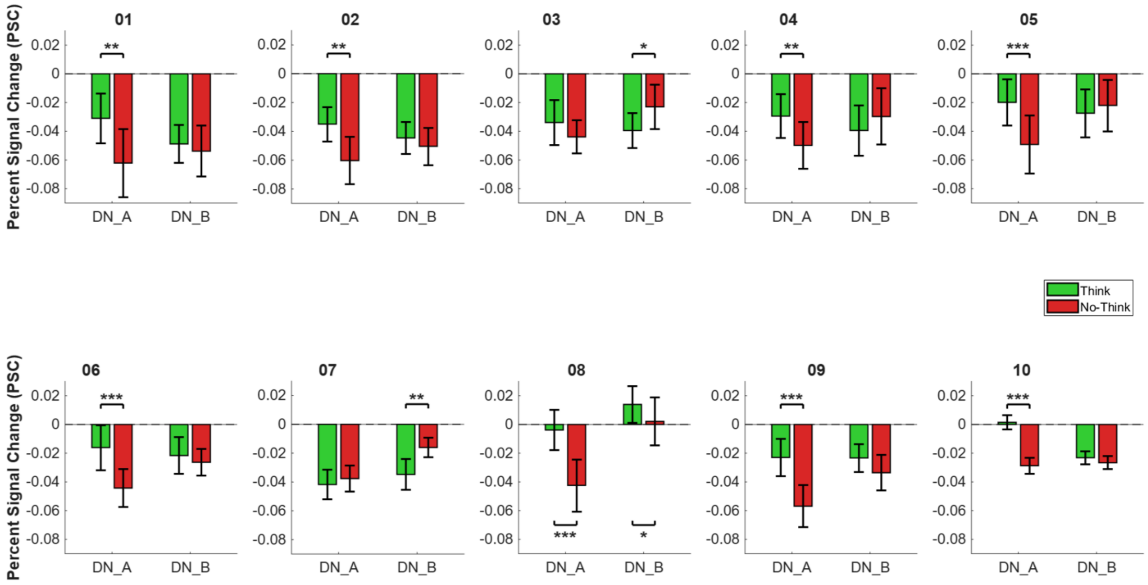


Figure 13: **Within-dataset analyses for Default Network subsystems.**
Each dataset is individually analyzed to observe trial-type effects on PSC in each Region Of Interest separately. Green bars represent PSC associated with Think condition, while red bars represent PSC associated with No-Think condition.
Sample size for each t-test can be extracted from demographic information of each dataset in table 1. Error bars show condition-specific 95% confidence intervals.
Statistical significance is shown as:$p < 0.05$ (*), $p < 0.01$ (**), $p < 0.001$ (***)

4.3 Correlation Between Deactivations

Correlations between deactivations were explored between Default Network subsystems and hippocampi. We wanted to explore these as an immediate proxy for behaviour similarity.

A statistically significant, strong positive correlation was found between differences in activation between PSC of Think and No-Think condition (Δ PSC) in hippocampi, and Δ PSC in Default Network A ($r = 0.74$, $p < 0.001$). This indicates that participants who exhibited a larger downregulation of activity in the hippocampi during No-Think trials in respect to the rest of the participants, also exhibited a proportionally larger downregulation in DN-A. This relationship accounts for roughly 55% of the shared variance ($R^2 = 0.548$).

Similarly, a significant positive correlation was found between hippocampi and DN-B PSC ($r = 0.68$, $p < 0.001$), explaining approximately 46% of the variance ($R^2 = 0.462$).

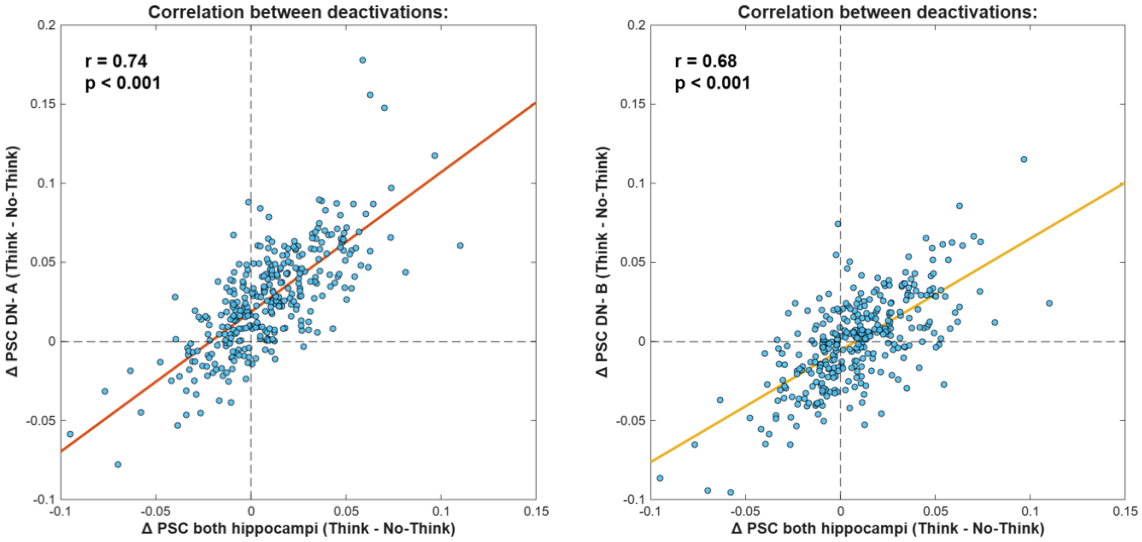


Figure 14: **Correlation between deactivations.**
Correlations between difference of activations between hippocampi and Default Network subsystems, on the left with Default Network A, on the right with Default Network B.
Each dot represents a subject, with total sample size $N = 330$.
Best fit lines are plotted for visualization purposes.

4.4 Qualitative Inspection Of Deactivations

The results described before all derive from a methodological choice about Percent Signal Change (PSC) calculation. We employed "mean" PSC, that is, the time courses of all the voxels within a Region Of Interest were averaged together to create a mean time course for that ROI, from which activity measure was derived. This approach is valid for small, coherent, structural regions such as the hippocampal formation. Our hippocampal ROI were manually traced on individual structural scans to obtain the most accurate selection of voxels within the structural hippocampus, and functionally coherent anatomical structures should not have great differences of functional profiles between different voxels. What is instead critical is applying such measures for ROI such as Default Network A (DN-A) and Default Network B (DN-B) as derived from the Du15NET atlas [74].
First, both ROI have a considerable size, as Default Network B comprises 4104 3mm² voxels, and Default Network A comprises 2090 3mm² voxels (voxel count conducted on masks on MNI space before warping to subject-specific anatomy).
Second, they entail sparse functional areas and isolated voxels, which could have distinct activations in response to the TNT task.
Could it be that the deactivations that we find by looking at a mean time course are actually driven only by some areas, and the observed network subsystems do not show functional coherence?

To test this, we developed a second pipeline, in which contrast maps were derived in MNI space to observe group-level contrast of Think $>$ No-Think.
To have an account of coherent deactivation in each subsystem, we parcellated the Default Network masks in their major functional areas, and retrieved an anatomical or functional label for each area. If no lateralization is described, it means that areas are along midline regions.
For DN-A, the major functional areas were: bilateral superior middle temporal gyrus (sMTG), bilateral posterior inferior parietal lobule (pIPL), bilateral parahippocampal cortex (PHC), ventral precuneus, posterior cingulate cortex (PPC), ventro-medial prefrontal cortex (vmPFC), a bilateral cluster in dorsal prefrontal cortex (dlPFC), left lateral rostral prefrontal cortex (lRPFC) (also named frontopolar cortex).

For DN-B, major areas described were: bilateral middle temporal gyrus (MTG), bilateral anterior inferior frontal gyrus (aIFG), bilateral lateral orbitofrontal cortex (lOFC), bilateral temporo-parietal junction (TPJ), dorsal and anterior medial prefrontal cortex (mPFC), posterior cingulate cortex (PPC), dorsal cingulate cortex (DCC), left dorso-lateral prefrontal cortex (dlPFC).

Since second-level analyses were performed on a sample size of 330 participants, many voxels in the $T > NT$ contrast exhibit a statistically significant effect, even if we applied rigid thresholding values (Family-Wise Error (FWE) corrected threshold of $p < 0.01$, with a cluster forming threshold of k=10). We decided to interpret this contrast via a Cohen's d map thresholded at a small effect size of 0.2. Each voxel in Cohen's d map surviving this threshold had also a statistically significant effect on the p-value map.
Sagittal slices were taken at various positions to observe Default Network subsystem areas. Since the contrast is on Think condition > No-Think condition, no information is provided if voxels were more active than baseline during the Think condition, or were having a downregulation during No-Think condition. For this reason, interpretation of contrast effect size is reserved to Default Network subsystems areas, to again corroborate that deactivation found on the mean signal of the subsystems are not driven by only a subset of areas.
At most lateral slice, at X = +60, we can observe DN-A's right superior MTG showing a few voxels with a small effect size, while DN-B areas do not show any effect. At X = ± 50, we see DN-B's areas bilateral TPJ, bilateral aIFG and part of bilateral MTG showing no effect size, while part of DN-A's right pIPL showing some voxels displaying effect. Only a few voxels are highlighted in the left superior MTG.
At X = -36 we can see left pIPL showing a medium-to-high effect size, while DN-B's cluster in dlPFC and remaining left MTG again do not show any effect. At X = ± 30 we can observe Hippocampal formation and DN-A's parahippocampal cortex showing medium effect sizes, while DN-B's lOFC doesn't show effect, nor does the rest of DN-B's left dlPFC cluster.
At X = -18 and X= +26 we see left and right clusters of DN-A in the dlPFC with clear effects, but left lRPFC defies expectancy as it does not show task modulation.
At midline slices (X = ± 4) we see ubiquitous effects in many regions. We have medium-to-large effect sizes in the visual cortex, as described elsewhere [96]. We see medium effect sizes in ventral Precuneus and in voxels in PCC, both of DN-A and DN-B. Dorsal Cingulate Cortex seems to not exhibit any effect. Most dorsal and anterior mPFC has no activation voxels, while a few voxels of DN-A's ventral mPFC show small effect sizes.

We have recapitulated all major areas parcellated from each Default Network subsystem, showing how only two areas of DN-A do not show effect for the $T > NT$ contrast, which are the small clusters in superior MTG and the left lRPFC. On the contrary, all clusters of DN-B except PCC do not exhibit any effect.
Taken together, these results suggest that DN-A overall deactivation is driven by all areas, functioning as a coherent network.

They also corroborate DN-B's lack of modulation during the TNT task, showing how almost no area within the subsystem displays a $T > NT$ effect, again confirming the functional coherence of the subsystem.
It is striking to observe the aforementioned dissociation in group-level task contrast: these accounts are noisy, as we can see from the effect-size map, and anatomical and functional variability can alter spatial precision of results.

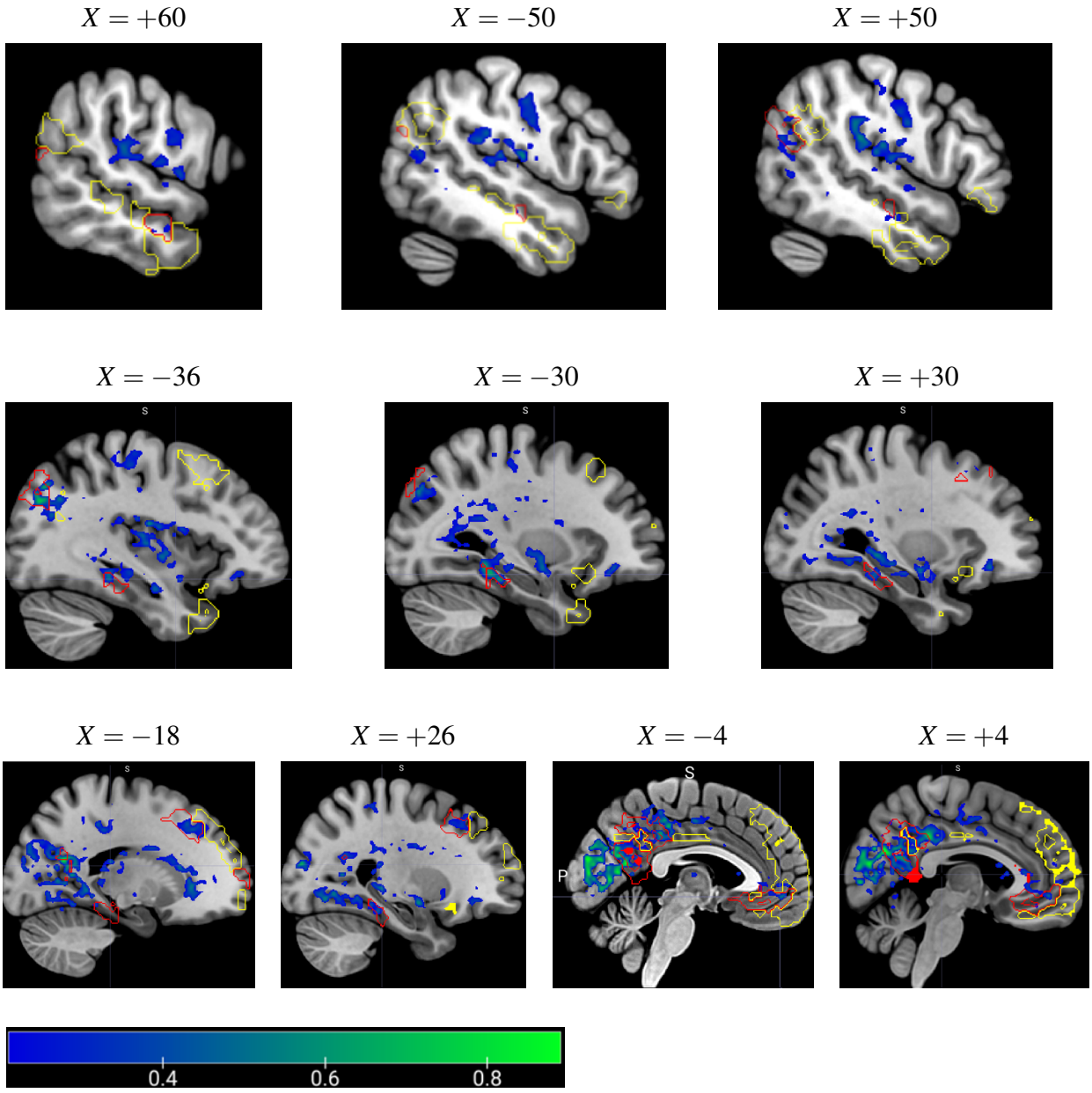


Figure 15: **Exploratory whole-brain analyses.**
Blue-to-green voxels represent effect sizes for the Think>No-Think contrast. Red outline marks edges of Default Network A areas, yellow outline for Default Network B.

5 Discussion

The results presented in this thesis shed light on neural correlates of the Retrieval Stopping process. Past literature had focused on the right hippocampus as an *a-priori* Region Of Interest (ROI) to investigate downregulation related to the No-Think condition, so works in subsequent years had focused only on the lateralized structure.

The novelty of the present approach is conducting analyses in subjects' native space with highly precise masks for ROIs, providing results with high accuracy and spatial precision. Corroborating early evidence, in this mega-analysis pooling data from 330 individuals, we found that both the right and left hippocampus, as manually traced on structural scans for each subject, showed reduced activity during Retrieval Stopping in respect to active retrieval, across a variety of experimental items, scanning sequences and laboratories.
We calculated an effect size for each structure independently, and found that the right hippocampus showed a reliable larger effect size than the left hippocampus (Cohen's $d = 0.343$ and Cohen's $d = 0.273$ for right and left respectively), finally confirming a long-held theory of the laboratory.

While the contrast Think $>$ No-Think can elucidate various mechanisms, it does not provide information if activity during Think condition is increased in respect to baseline activity, or No-Think condition is associated with a downregulation. We calculated the measure of Negative BOLD Responses (NBRs) as a proxy for downregulation: both the right and left hippocampi exhibited NBRs associated with No-Think trials, coherent with other work [11, 22]. In light of results obtained elsewhere [22, 51, 53], we can interpret NBRs as indicating the hippocampus being properly inhibited, as the structures are being targeted by an inhibitory control mechanism, centered around right anterior dorso-lateral prefrontal cortex and right ventro-lateral prefrontal cortex [97].
Contrary to past work [11], we do not observe Percent Signal Change associated with Think trials to be positive in respect to baseline. Observing activation above baseline in the hippocampus is a remarkably hard inquiry, even with memory heavily-dependent tasks, due to hippocampal high activity even during rest or null trials, often used as data points for defining baseline [98].

A previous study pointed to a possible modulation, as activity was greater during Think trials than during No-Think trials, of areas historically associated with the Default Network during the TNT task, subsequently described as target of a "broader deactivation pattern" [11, 27]. To corroborate this hypothesis, we implemented the most recent parcellation of human brain resting state networks [74], which by taking advantage of precision

imaging differentiates two networks within the Default Network.
We did find that Retrieval Stopping caused a reduction of activity in respect to active retrieval in Default Network, but found that this modulation was exclusively driven by reduced activity in Default Network A (DN-A), while the other network of interest, termed Default Network B (DN-B), was not modulated by the TNT task.
We found a significant effect of dataset type on the interaction between trial-type effect and ROI effect: we conducted analyses within each datasets, and found that most dataset recapitulated overall effect, i.e. reduced activity in DN-A between Think and No-Think and no modulation of DN-B.

To investigate downregulation of the networks during Retrieval Stopping, we calculated NBRs. In contrast with hippocampal results, we demonstrated how both Retrieval Stopping attempts and active retrieval attempts produced NBRs, both in DN-A and DN-B. Given the documented behaviour of Default Network and its functional profile, these NBRs should be interpreted with caution.
The baseline from which NBRs were calculated were data points scattered around the scanning session: while intuition may tell us that participants were at rest during these moments, left without any behavioural task to do nor cognitive operation to follow, we know that "brain at rest does not rest". The brain engages in unconstrained cognitive activity during task-free moments: spontaneous cognition and autobiographical planning, stimulus-independent thought and rumination accompany us in almost every spare waking moment.
One way to interpret the observed NBRs in both Think and No-Think conditions, in both networks, is that they are caused by the difference between spare moments and complying with the paradigm: performing any task with clear instructions and goals could in theory constrain mental activity. Even during Think trials, thinking and retrieving a single target memory that was extensively studied can constrain the mind-wandering and day-dreaming that we experience during spare moments. Importantly, Default Network has been historically related to this day-dreaming activity [76], and early evidence showed how allocating attention away from this unconstrained internal cognition was related to downregulation of the various cortical areas associated with the network [99].

Reduced activity observed in both DN-A and DN-B in respect to baseline could thus be explained by the difference between attending to the task and unrestrained mental activity. Indeed, researchers argue how, especially for memory tasks, the task-free resting state is not to be considered a stable baseline [98]. This problem is a direct consequence of the main setback of BOLD fMRI, which is its contrastive nature. The methodology does not provide a true baseline-signal level, and it is not possible to determine how much activity is associated in certain brain regions with a condition alone, but contrast between conditions are required.

While this interpretation reduces the explaining power of the NBRs observed, it does not affect the importance of the measured reduction in activity during No-Think trials in respect to Think trials in DN-A, together with no observed modulation of DN-B.

These results are confirmed also by the ROIs behaving as functionally coherent networks, as most areas in DN-A show an effect for $T > NT$, while almost no DN-B's area contains voxels in which the contrast has any small effect.

We argue that the DN-A modulation is a result of Retrieval Stopping attempt, that is, a downregulation initiated from the inhibitory control mechanism centered in the right anterior dorso-lateral prefrontal cortex and right ventro-lateral prefrontal cortex.

DN-A role in episodic construction has been extensively studied: it has been demonstrated to be implicated in recollection of prior information and episodic retrieval, and supports spatio-contextual memory, playing an overarching role in memory-based construction of events [66, 100–102]. Furthermore, a recent high-resolution investigation has provided evidence for connectivity differences between DN-A and DN-B to areas in medial temporal lobe (MTL), with DN-A possibly representing cortical areas subserving memory-related functions thanks to interactions with certain MTL structures [103]. On the contrary, DN-B seems to have a role in reflecting on the mental states of other people and also one's own mental states: it is selectively engaged when participants assess social information [72, 103]. DN-B's distinct role in social cognition does not suggest any involvement in memory tasks such as the TNT task, and we indeed see no modulation of this network in our data.

We thus propose that DN-A experiences the same effect as hippocampi as they work together to enable memory retrieval. Inhibitory control seeks the specific cancellation of retrieval activity within both hippocampus and the DN-A, while sparing the uninvolved DN-B.

Our proposal is also supported by the strong correlation we saw in individuals between downregulation effect in hippocampi and downregulation effect in DN-A. Individuals who achieved great reduction of hippocampal activity during No-Think trials relative to Think trials also managed great reduction in activity in DN-A, whereas those who showed little to no reduction in the hippocampi similarly failed to downregulate DN-A. This provides evidence for functional co-modulation, pointing to a unique mechanism affecting both systems with a similar amplitude. Unexpectedly, a significant positive correlation was also found between difference in activity of hippocampi and difference in activity of DN-B, weaker relative to DN-A's one. This relationship could reflect intra-subject Default Network subsystems covariance, rather than functional similarity between Hippocampi and DN B, but further data would be needed to test this.

As said before, part of the research history (and historical definition) of the Default Network is its prominent downregulation during complex, attention-demanding tasks. A counter-interpretation of our results could be that DN-A is modulated because No-Think trials require more attention, or are simply "harder", than retrieving an extensively studied associate given its cue.
The effect we found for Think>No-Think could thus be explained solely by the trial difference in cognitive resources demanded, and not by an underlying cognitive mechanism such as Retrieval Stopping.
While this counter-argument may have some basis, there are a few fallacies in this view.

Firstly, even if past literature searched for deactivations in canonical Default Network before precision parcellation, no paper seems to provide evidence for a distinct pattern of task-induced deactivation in specific areas of the network. Beyond the functional specificity and different connectivity, DN-A and DN-B still share identical deactivations during externally-focused tasks [74]. There's thus no reason to assume that Retrieval Stopping would cause decreased activations only in DN-A, while DN-B remains unaffected.

Second, while originally it was thought that Default Network deactivations were caused by reallocation of processing resources and that deactivations scaled with task overall "difficulty" [104], we now know that this is not exact. Rather, it seems that DN deactivations are not related to the difficulty of the task we are attending, but by its quality. External-attention demanding tasks, such as perceptual assignments on visual stimuli, greatly deactivate the network [105], whilst tasks with similar perceived "hardness" or complexity that require attention focused on our internal world do not show equal DN downregulations [106] [107]. There's no reason to assume that the Retrieval Stopping process would require intense focused attention on the external world, while active retrieval wouldn't.

A direct way to falsify the proposed counter-argument would be to analyze results from an alternative way of conducting the TNT paradigm, introduced by Roland Benoit [48]. In this paradigm, participants are instructed to avoid retrieving the target memory during No-Think trials by retrieving any alternative thought, strategy termed "thought substitution". The paper demonstrates how this strategy does not cause the same memory disruption effects as direct Retrieval Stopping, and does not result in hippocampus inhibition [48]. In the future, we could thus assess DN NBRs in this paradigm: during "thought substitution" trials participants are instructed to not think about the target, the trial engages cognitive resources, and there's no reason to suppose it is more attention-demanding than direct suppression instructions.
If in this hypothetical DN-A activity wouldn't be modulated by No-Think or Think trials

(as hippocampal activity is not modulated), it would provide a proof of evidence that the effect we see in our results can't be explained by resource-demanding attention-related downregulation, but it is due to a selective downregulation of DN-A activity, while sparing DN-B.

Our thesis, other than bringing clarifications on the cortical correlates of Retrieval Stopping, also brings evidence, on top of the one accumulated by Buckner's lab, Rodrigo Braga and Andrews-Hanna, that the canonical, coherent, integrated Default Network is a myth [66, 67, 71, 74, 103, 108]. Evidence shows undeniably how two networks within the canonical network, with their own distinct functional profile, distinct anatomical locations, distinct connections with other brain areas, should be considered as separate entities. While in this work the term "subsystem" was extensively used, we expect that the two networks will be regarded in the near future as entirely distinct, with no overall system formed by their union.
Past low-resolution and group-averaged data must not obscure what recent evidence tells us. Furthermore, the impression that there is a coherent monolithic network may have been reinforced by the phenomenon of task-induced deactivation, which, as we now know, expands across various association areas and different networks [74].

While no contrast was explored between Percent Signal Change results in native space and in standard space, our results argue for performing analyses in native space, even by using standard atlas-based Region Of Interest warped to each subject's brain. While few studies have used this technique and advocated for it [109–111], this is still not the norm in neuroimaging literature. The results presented couldn't have been obtained with standard atlas-based analyses: since DN-A and DN-B, as defined in our parcellation of choice [74], are closely juxtaposed and interdigitated, warping transformations and smoothing would have heavily affected final results. We advocate for native-space analyses, especially given that inverse parameters $(-y)$ to be followed to conduct warping from atlas of choice to native space are already calculated in standard pipelines, so no additional step is required. However, we understand how greater computational power needed during analyses can be a hindrance to researchers.

The results presented here have operative implications in the context of mental health. Indeed, many mental health conditions, such as schizophrenia, depression, anxiety disorders and obsessive-compulsive disorder are hallmarked by symptoms such as hallucinations, intrusive thoughts, ruminations about the past, fearful future scenarios or compulsive thinking [112].
The theoretical framework adopted in the thesis posits that a domain-general inhibitory control mechanism mediated by the right anterior dorso-lateral prefrontal cortex and the

ventro-lateral prefrontal cortex can dynamically target the medial temporal lobe to achieve thought suppression [27].
Thus, whether intrusions concern the past or future, the real or hypothetical, general thoughts or specific events, the fronto-temporal pathway may stop their retrieval. While most frameworks on psychiatric pathology focus on altered executive functions and their prefrontal basis, our theoretical framework emphasizes the target of that control: namely, the hippocampus.
It has been demonstrated that concentrations of hippocampal GABA in humans, as assayed by magnetic resonance spectroscopy, predict the strength of inhibitory coupling between the prefrontal cortex and the hippocampus, and subsequent ability of Retrieval Stopping [22]. A plethora of studies have demonstrated how hippocampal GABAergic activity is diminished in schizophrenia [113, 114], and hippocampal hyperactivity is present in depression [115, 116], anxiety [117], post-traumatic stress disorder [118], phenomenon aptly named "hippocampal disinhibition".
Indeed, studies have found that a range of psychological disorders are associated with impaired suppression-induced forgetting, as a proxy for inhibitory control on internal thoughts, relative to control samples [119].
Thus, a disrupted inhibitory pathway between prefrontal inhibitory regions and MTL regions allows us to explain transdiagnostic origins of intrusive thinking, resulting in core symptoms of various mental health disorders.
Our results point to the downregulation of Default Network A areas as another consequence of inhibitory control on thoughts. Supported by research showing that neural correlates of self-talk and internal cognition closely match Default Network areas [76, 120, 121], they provide a new framework to explore transdiagnostic markers for mental health. Indeed, an influential review found that the Default Network was hyperactive across mental illnesses, and showed reduced segregation measures with fronto-parietal networks [122]. We thus propose that reduced hippocampal GABA hinders the ability to suppress intrusive thoughts, disinhibiting both the hippocampus and broader Default Network, causing persistent intrusive thoughts and a tendency to focus attention inwardly [123, 124].

Many research directions can stem from the results presented here.
First of all, further analyses should be conducted to explain the numerous cortical voxels showing statistically significant effects for the Think>No-Think contrast in MNI space found in this mega-analysis. Many observed cortical regions sensitive to the contrast seem to not have been described before in the TNT literature.
Second, techniques should be implemented to test if Default Network A areas are properly inhibited, or if downregulation observed is driven by BOLD signal artifacts. While some

research is pointing at NBRs as being predictive for neural tissue inhibition [125], we are still cautious to interpret our results as inhibition.

If, as we suspect, an inhibition process is in place, further work should resolve if inhibition pathways target in parallel hippocampus and DN-A, or alternatively, if DN-A activity is driven by hippocampus, so that inhibiting this latter structure results in lower activity in the network. Another research direction, easily implementable, would be to assess functional connectivity between hippocampus and Default Network subsystems, even if significance of functional connectivity during task is still being discussed [126].

Finally, we could investigate if in mental conditions hallmarked by intrusive thoughts there's a diminished downregulation of Default Network A during Retrieval Stopping, paralleling the observed impairment in inhibiting hippocampus by patients attempting the Think/No-Think task [127, 128].

6 Conclusions

Findings presented here shed light on the neural correlates of Retrieval Stopping thanks to functional Magnetic Resonance Imaging scanning during execution of the Think/No-Think task.

With a big sample size and procedures conducted in native space, this mega-analysis provides evidence for reduced activity in both the right and left hippocampus during Retrieval Stopping in respect to active retrieval. Furthermore, when activity is compared between conditions and during unmodeled data points (signal baseline), Retrieval Stopping causes a negative deflection of the signal, i.e. a Negative Bold Response (NBR), while in our data active retrieval does not cause a positive deflection. We found a bigger effect size in the right hippocampus relative to the left hippocampus, which can explain why past work has focused on the right lateralized structure.

We found that Default Network A (DN-A), as defined by a recent parcellation of the human brain in ten resting-state networks with high precision, recapitulates hippocampal modulation. DN-A shows reduced activity during Retrieval Stopping relative to active retrieval, in both its overall mean signal and signal from each area that composes the network. DN-A activity is substantially lower than baseline during Retrieval Stopping, but exhibits an NBR also during Think trials.

Importantly, Default Network B is not modulated by the Think/No-Think task. It exhibits NBRs for both Retrieval Stopping and active retrieval, with no significant statistical difference. This is corroborated by looking at single areas composing the network as well.

We furthermore found a correlation between downregulation in hippocampi and downregulation in DN-A. While cortical regions may share nuisance variance that could explain this correlation, we find that the correlation coefficient was higher for DN-A relative to DN-B.

These results, taken together, corroborate functional specialization between DN-A and DN-B.

We argue that both the right and left hippocampus and Default Network A cooperate together to enable conscious memory retrieval.

We argue that a domain-general inhibitory control mechanism seeks the cancellation of retrieval activity during Retrieval Stopping, and thus disrupts activity of both hippocampus and DN-A while sparing activity of the uninvolved Default Network B.

Bibliography

[1] Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica*, 30, 412–431. (Original work published 1868).

[2] Ebbinghaus, H. (1885). *Über das gedächtnis: untersuchungen zur experimentellen psychologie*. Duncker & Humblot.

[3] Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological review*, 55(4), 189.

[4] Bari, A., & Robbins, T. (2013). Inhibition and impulsivity: behavioral and neural basis of response control.. *Progress in neurobiology*, 108, 44-79 .

[5] Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature*, 410(6826), 366-369.

[6] Nardo, D., & Anderson, M. C. (2024). Everything you ever wanted to know about the Think/No-Think task, but forgot to ask. *Behavior Research Methods*, 56(4), 3831-3860.

[7] Anderson, M. C., Ochsner, K. N., Kuhl, B., Cooper, J., Robertson, E., Gabrieli, S. W., Glover, G. H., & Gabrieli, J. D. E. (2004). Neural Systems Underlying the Suppression of Unwanted Memories. *Science*, 303(5655), 232-235.

[8] Garavan H, Ross TJ, Murphy K, Roche RA, Stein EA. Dissociable executive functions in the dynamic control of behavior: inhibition, error detection, and correction. *Neuroimage*. 2002 Dec;17(4):1820-9.

[9] Aron, A. R., Robbins, T. W. & Poldrack, R. A. Inhibition and the right inferior frontal cortex: one decade on. *Trends in Cognitive Sciences* 18, 177–185 (2014).

[10] Apšvalka, D., Ferreira, C. S., Schmitz, T. W., Rowe, J. B., & Anderson, M. C. (2022). Dynamic targeting enables domain-general inhibitory control over action and thought by the prefrontal cortex. *Nature communications*, 13(1), 274.

[11] Levy, B. J. & Anderson, M. C. (2012). Purging of memories from conscious awareness tracked in the human brain. *The Journal of Neuroscience* 32, 16785–16794.

[12] Waldhauser, G. T., Braun, V. & Hanslmayr, S. (2016). Episodic memory retrieval functionally relies on very rapid reactivation of sensory information. *The Journal of Neuroscience* 36, 251–260.

[13] Gagnepain, P., Henson, R. N. & Anderson, M. C. (2014). Suppressing unwanted memories reduces their unconscious influence via targeted cortical inhibition. *Proceedings of the National Academy of Sciences* 111, E1310–E1319.

[14] Meyer, A. K., & Benoit, R. G. (2022). Suppression weakens unwanted memories via a sustained reduction of neural reactivation. *elife*, 11, e71309.

[15] Gagnepain, P., Hulbert, J. & Anderson, M. C. (2017). Parallel regulation of memory and emotion supports the suppression of intrusive memories. *The Journal of Neuroscience* 37, 6423–6441.

[16] White, E. L., & Keller, A. (1989). *Cortical circuits: synaptic organization of the cerebral cortex–structure, function, and theory*. Boston: Birkhäuser.

[17] Anderson, M. C., Bunce, J. G., & Barbas, H. (2016). Prefrontal–hippocampal pathways underlying inhibitory control over memory. *Neurobiology of Learning and Memory*, 134, 145-161.

[18] Crespo-García, M., Wang, Y., Jiang, M., Anderson, M. C. & Lei, X. (2022). Anterior cingulate cortex signals the need to control intrusive thoughts during motivated forgetting. *The Journal of Neuroscience* 42, 4342–4359.

[19] Barbas, H., Ghashghaei, H., Dombrowski, S. M. & Rempel-Clower, N. L. Medial prefrontal cortices are unified by common connections with superior temporal cortices and distinguished by input from memory-related areas in the rhesus monkey. *Journal of Comparative Neurology* 410, 343–367 (1999).

[20] Fernandez, F., Via, G., Canavier, C., & White, J. A. (2022). Kinetics and Connectivity Properties of Parvalbumin- and Somatostatin-Positive Inhibition in Layer 2/3 Medial Entorhinal Cortex. *eNeuro*, 9.

[21] Andrianova, L. et al. (2021). Hippocampal CA1 pyramidal cells do not receive monosynaptic input from thalamic nucleus reuniens. Preprint at *bioRxiv*.

[22] Schmitz, T. W., Correia, M. M., Ferreira, C. S., Prescot, A. P. & Anderson, M. C. (2017). Hippocampal GABA enables inhibitory control over unwanted thoughts. *Nature communications* 8(1), 1311.

[23] Hangya, B., Borhegyi, Z., Szilágyi, N., Freund, T. F., & Varga, V. (2009). GABAergic neurons of the medial septum lead the hippocampal network during theta activity. *The Journal of neuroscience*, 29(25), 8094-8102.

[24] Tóth, K., Freund, T. F., & Miles, R. (1997). Disinhibition of rat hippocampal pyramidal cells by GABAergic afferents from the septum. *The Journal of physiology*, 500(2), 463-474.

[25] Schnider, A. (2003). Spontaneous confabulation and the adaptation of thought to ongoing reality. *Nature Reviews Neuroscience* 4, 662–671.

[26] Wessel, J. R., & Aron, A. R. (2017). On the globality of motor suppression: unexpected events and their influence on behavior and cognition. *Neuron*, 93(2), 259-280.

[27] Anderson, M. C., Crespo-Garcia, M., & Subbulakshmi, S. (2025). Brain mechanisms underlying the inhibitory control of thought. *Nature Reviews Neuroscience*, 26(7), 415-437.

[28] Braak, H. (1974). On the structure of the human archicortex: I. The cornu ammonis a Golgi and pigmentarchitectonic study. *Cell and tissue research*, 152(3), 349-383.

[29] Witter, M., Wouterlood, F., Naber, P. A., & Van Haeften, T. (2000). Anatomical Organization of the Parahippocampal-Hippocampal Network. *Annals of the New York Academy of Sciences*, 911.

[30] Deshmukh, S. S., & Knierim, J. J. (2012). Hippocampus. *WIREs Cognitive Science*, 3(2), 231-251.

[31] Gross, C. G. (1993). Huxley versus Owen: the hippocampus minor and evolution. *Trends in Neurosciences*, 16(12), 493-498.

[32] Nadel L, O'Keefe J, Black A (1975). "Slam on the brakes: a critique of Altman, Brunner, and Bayer's response-inhibition model of hippocampal function". *Behavioral Biology*. 14 (2): 151–162.

[33] Scoville, W. B., & Milner, B. (1957). LOSS OF RECENT MEMORY AFTER BILATERAL HIPPOCAMPAL LESIONS. *Journal of Neurology, Neurosurgery & Psychiatry*, 20(1), 11–21.

[34] Teyler, T., & DiScenna, P. (1986). The hippocampal memory indexing theory. *Behavioral neuroscience*, 100 2, 147-54.

[35] McClelland, J. L., McNaughton, B., & O'Reilly, R. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological review*, 102 3, 419-457.

[36] Takashima, A., Nieuwenhuis, I., Jensen, O., Talamini, L., Rijpkema, M., & Fernández, G. (2009). Shift from Hippocampal to Neocortical Centered Retrieval Network with Consolidation. *The Journal of Neuroscience*, 29, 10087 - 10093.

[37] Squire, L. R., Genzel, L., Wixted, J., & Morris, R. G. M. (2015). Memory consolidation. *Cold Spring Harbor perspectives in biology*, 7 8, a021766.

[38] Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology*, 16(6), 693-700.

[39] Iwata, T., Yanagisawa, T., Ikegaya, Y., Smallwood, J., Fukuma, R., Oshino, S., Tani, N., Khoo, H.-M., & Kishima, H. (2024). Hippocampal sharp-wave ripples correlate with periods of naturally occurring self-generated thoughts in humans. *Nature Communications*, 15.

[40] McCormick, C., Rosenthal, C. R., Miller, T., & Maguire, E. (2018). Mind-Wandering in People with Hippocampal Damage. *The Journal of Neuroscience*, 38, 2745 - 2754.

[41] Morris, R. G. M., Garrud, P., Rawlins, J., & O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature*, 297, 681-683.

[42] O'Keefe, J., & Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research*, 34(1), 171-175.

[43] Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., & Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052), 801-806.

[44] Solstad, T., Boccara, C. N., Kropff, E., Moser, M. B., & Moser, E. I. (2008). Representation of geometric borders in the entorhinal cortex. *Science*, 322(5909), 1865–1868.

[45] Taube, J., Muller, R. U., & Ranck, J. B. (1990). Head-direction cells recorded from the postsubiculum in freely moving rats. I. Description and quantitative analysis. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 10(2), 420–435.

[46] Buzsáki, G., & Moser, E. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nature neuroscience*, 16, 130 - 138.

[47] Eichenbaum, H. (2017). The role of the hippocampus in navigation is memory. *Journal of neurophysiology*, 117 4, 1785-1796.

[48] Benoit, R. G. & Anderson, M. C. (2012). Opposing mechanisms support the voluntary forgetting of unwanted memories. *Neuron* 76, 450–460.

[49] Benoit, R. G., Hulbert, J. C., Huddleston, E. & Anderson, M. C. (2015). Adaptive top-down suppression of hippocampal activity and the purging of intrusive memories from consciousness. *Journal of Cognitive Neuroscience* 27, 96–111.

[50] Benoit, R. G., Davies, D. J. & Anderson, M. C. (2016). Reducing future fears by suppressing the brain mechanisms underlying episodic simulation. *Proceedings of the National Academy of Sciences* 113, E8492–E8501.

[51] Crespo-Garcia, M., Wang, Y., Jiang, M., Anderson, M. C. & Lei, X. (2022). Anterior cingulate cortex signals the need to control intrusive thoughts during motivated forgetting. *The Journal of Neuroscience* 42, 4342–4359.

[52] Suarez, A., Valdés-Hernández, P., Bernal, B., Dunoyer, C., Khoo, H.-M., Bosch-Bayard, J., & Riera, J. (2021). Identification of Negative BOLD Responses in Epilepsy Using Windkessel Models. *Frontiers in Neurology*, 12.

[53] Hulbert, J. C., Henson, R. N. & Anderson, M. C. (2016). Inducing amnesia through systemic suppression. *Nature communications* 7, 11003.

[54] Waldhauser, G. T., Bäuml, K.-H. T., & Hanslmayr, S. (2014). Brain Oscillations Mediate Successful Suppression of Unwanted Memories. *Cerebral Cortex*, 25(11), 4180-4190.

[55] Callard F and Margulies DS (2014) What we talk about when we talk about the default mode network. *Frontiers in human neuroscience* 8:619.

[56] Ingvar, D.H. (1979), "Hyperfrontal" distribution of the cerebral grey matter flow in resting wakefulness; on the functional anatomy of the conscious state. *Acta Neurologica Scandinavica*, 60: 12-25.

[57] Shulman, G.L., Fiez, J.A., Corbetta, M., Buckner, R.L., Miezin, F.M., Raichle, M.E., and Petersen, S.E. (1997). Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *Journal of cognitive neuroscience* 9, 648–663.

[58] Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., and Shulman, G.L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences* 98, 676–682.

[59] Greicius, M. D., & Menon, V. (2004). Default-Mode Activity during a Passive Sensory Task: Uncoupled from Deactivation but Impacting Activation. *Journal of Cognitive Neuroscience*, 16(9), 1484-1492.

[60] Greicius, M. D., Krasnow, B., Reiss, A. L., and Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences* 100, 253–258.

[61] Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: An fMRI investigation of the resting-state default mode of brain function hypothesis. *Human Brain Mapping*, 26.

[62] Damoiseaux, J., Rombouts, S. A., Barkhof, F., Scheltens, P., Stam, C., Smith, S. M., & Beckmann, C. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences*, 103, 13848 - 13853.

[63] Gusnard, D., Akbudak, E., Shulman, G. L., & Raichle, M. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proceedings of the National Academy of Sciences*, 98, 4259 - 4264.

[64] Spreng, R., Mar, R., & Kim, A. S. N. (2009). The Common Neural Basis of Autobiographical Memory, Prospection, Navigation, Theory of Mind, and the Default Mode: A Quantitative Meta-analysis. *Journal of Cognitive Neuroscience*, 21, 489-510.

[65] Alves, P.N., Foulon, C., Karolis, V. et al. (2019). An improved neuroanatomical model of the default-mode network reconciles previous neuroimaging and neuropathological findings. *Communications biology* 2, 370.

[66] Buckner, R. L., & DiNicola, L. M. (2019). The brain's default network: updated anatomy, physiology and evolving insights. *Nature Reviews Neuroscience*, 20(10), 593-608.

[67] Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-Anatomic Fractionation of the Brain's Default Network. *Neuron*, 65(4), 550-562.

[68] Campbell, K. L., Grigg, O., Saverino, C., Churchill, N., & Grady, C. L. (2013). Age differences in the intrinsic functional connectivity of default network subsystems. *Frontiers in aging neuroscience*, 5, 73.

[69] Hayes, S. M., Salat, D. H., & Verfaellie, M. (2012). Default network connectivity in medial temporal lobe amnesia. *The Journal of neuroscience*, 32(42), 14622-14629a.

[70] Wen, T., Mitchell, D. J., & Duncan, J. (2020). The Functional Convergence and Heterogeneity of Social, Episodic, and Self-Referential Thought in the Default Mode Network. *Cerebral Cortex*, 30(11), 5915-5929.

[71] Braga, R. M., & Buckner, R. L. (2017). Parallel interdigitated distributed networks within the individual estimated by intrinsic functional connectivity. *Neuron*, 95(2), 457-471.

[72] DiNicola, L. M., Braga, R. M., & Buckner, R. (2019). Parallel Distributed Networks Dissociate Episodic and Social Functions Within the Individual. *bioRxiv*.

[73] Braga, R. M., DiNicola, L. M., Becker, H. C., & Buckner, R. (2019). Situating the left-lateralized language network in the broader organization of multiple specialized large-scale distributed networks. *Journal of Neurophysiology*, 124, 1415 - 1448.

[74] Du, J., DiNicola, L. M., Angeli, P. A., Saadon-Grosman, N., Sun, W., Kaiser, S., Ladopoulou, J., Xue, A., Yeo, B. T. T., Eldaief, M. C., & Buckner, R. L. (2024). Organization of the human cerebral cortex estimated within individuals: networks, global topography, and function. *Journal of Neurophysiology*, 131(6), 1014-1082.

[75] Weissman, D., Roberts, K., Visscher, K. et al. (2006). The neural bases of momentary lapses in attention. *Nature neuroscience* 9, 971–978.

[76] Menon, V. (2023). 20 years of the default mode network: a review and synthesis. *Neuron* 111, 2469–2487.

[77] James, W., Burkhardt, F., Bowers, F., & Skrupskelis, K. (1890). *The principles of psychology* (Vol. 1, No. 2, pp. 315-329). London: Macmillan.

[78] Kawashima, I., Hinuma, T., & Tanaka, S. C. (2023). Ecological momentary assessment of mind-wandering: meta-analysis and systematic review. *Scientific Reports*, 13.

[79] Christoff, K., Gordon, A.M., Smallwood, J., Smith, R., and Schooler, J.W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences* 106, 8719–8724.

[80] Du, J.-N., Tripathi, V., Elliott, M. L., Ladopoulou, J., Sun, W., Eldaief, M., & Buckner, R. (2025). Within-individual precision mapping of brain networks exclusively using task data. *Neuron*, 113, 4069 - 4083.e6.

[81] McRobbie, D. W. (2007). *MRI from Picture to Proton*. United Kingdom: Cambridge University Press.

[82] Logothetis, N., Pauls, J., Augath, M., Trinath, T., & Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature*, 412, 150-157.

[83] Poldrack, R. A. (2012). The future of fMRI in cognitive neuroscience. *NeuroImage*, 62(2), 1216-1220.

[84] Obriot, J., Amor, Z., Serger, M., Stirnberg, R., Ehses, P., Riedel, M., ... & Boulant, N. (2026). Human fMRI at 11.7 T: Assessing feasibility, stability, and reliability on the Iseult scanner. *Imaging neuroscience*, 4, IMAG-a.

[85] Costafreda, S. (2009). Pooling fMRI data: meta-analysis, mega-analysis and multicenter studies. *Frontiers in Neuroinformatics*, 3.

[86] Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of Functional MRI Data Analysis*. Cambridge University Press.

[87] Parker, D. B., & Razlighi, Q. (2019). The Benefit of Slice Timing Correction in Common fMRI Preprocessing Pipelines. *Frontiers in Neuroscience*, 13.

[88] Wellcome Centre for Human Neuroimaging. (2014). *Statistical parametric mapping: SPM12 manual* (Version 12).

[89] Poldrack, R.A., 2007. Region of interest analysis for fMRI. *Social cognitive and affective neuroscience* 2, 67–70.

[90] Pruessner, J. C., Li, L. M., Serles, W., Pruessner, M., Collins, D. L., Kabani, N., ... & Evans, A. C. (2000). Volumetry of hippocampus and amygdala with high-resolution MRI and three-dimensional analysis software: minimizing the discrepancies between laboratories. *Cerebral cortex*, 10(4), 433-442.

[91] Rorden C. (2025). MRIcroGL: voxel-based visualization for neuroimaging. *Nature methods*, 22(8), 1613–1614.

[92] Kong, R., Li, J., Orban, C., Sabuncu, M. R., Liu, H., Schaefer, A., Sun, N., Zuo, X.-N., Holmes, A. J., Eickhoff, S. B., & Yeo, B. T. T. (2018). Spatial Topography of Individual-Specific Cortical Networks Predicts Human Cognition, Personality, and Emotion. *Cerebral Cortex*, 29(6), 2533-2551.

[93] Caballero-Gaudes, C., & Reynolds, R. (2016). Methods for cleaning the BOLD fMRI signal. *NeuroImage*, 154, 128 - 149.

[94] Murphy, K., Bodurka, J., & Bandettini, P. (2007). How long to scan? The relationship between fMRI temporal signal to noise ratio and necessary scan duration. *NeuroImage*, 34 2, 565-74.

[95] Knief, U., & Forstmeier, W. (2018). Violating the normality assumption may be the lesser of two evils. *Behavior Research Methods*, 53, 2576 - 2590.

[96] Yan, Y., Hulbert, J. C., Zhuang, K.-X., Liu, W., Wei, D.-T., Qiu, J., Anderson, M. C., & Yang, W. (2022). Reduced hippocampal-cortical connectivity during memory suppression predicts the ability to forget unwanted memories. *Cerebral Cortex (New York, NY)*, 33, 4189 - 4201.

[97] Wessel, J. R. & Anderson, M. C. Neural mechanisms of domain-general inhibitory control. *Trends in Cognitive Sciences* 28, 124–143 (2024).

[98] Stark, C. E., & Squire, L. R. (2001). When zero is not zero: the problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences*, 98(22), 12760-12766.

[99] McKiernan, K., D'Angelo, B. R., Kaufman, J., & Binder, J. (2006). Interrupting the "stream of consciousness": An fMRI investigation. *NeuroImage*, 29 4, 1185-91.

[100] Andrews-Hanna, J. R., Saxe, R., & Yarkoni, T. (2014). Contributions of episodic retrieval and mentalizing to autobiographical thought: Evidence from functional neuroimaging, resting-state connectivity, and fMRI meta-analyses. *NeuroImage*, 91, 324-335.

[101] DiNicola, L. M., Ariyo, O. I., & Buckner, R. (2022). Functional specialization of parallel distributed networks revealed by analysis of trial-to-trial variation in processing demands. *Journal of Neurophysiology*, 129, 17 - 40.

[102] Ritchey, M. (2026). What episodic memory reveals about the default mode network. *Current Opinion in Behavioral Sciences*, 69, 101646.

[103] Braga, R. M. (2025). The canonical default network comprises parallel distributed networks with distinct medial temporal lobe connections. *Current Opinion in Behavioral Sciences*, 66, 101610.

[104] McKiernan, K., Kaufman, J., Kucera-Thompson, J., & Binder, J. (2003). A Parametric Manipulation of Factors Affecting Task-induced Deactivation in Functional Neuroimaging. *Journal of Cognitive Neuroscience*, 15, 394-408.

[105] Singh, K. D., & Fawcett, I. P. (2008). Transient and linearly graded deactivation of the human default-mode network by a visual detection task. *NeuroImage*, 41 1, 100-12.

[106] Harrison, B., Davey, C., Savage, H. S., Jamieson, A., Leonards, C. A., Moffat, B. A., Glarin, R., & Steward, T. (2021). Dynamic subcortical modulators of human default mode network function. *Cerebral Cortex (New York, NY)*, 32, 4345 - 4355.

[107] Gilbert, S., Bird, G., Frith, C., & Burgess, P. (2012). Does "Task Difficulty" Explain "Task-Induced Deactivation?". *Frontiers in Psychology*, 3.

[108] Zamani, A., & Hadjiilieva, K. C. (2026). Default heterogeneity: from a singular construct to multiple default network (s). *Current Opinion in Behavioral Sciences*, 68, 101636.

[109] Gavrilescu, M., Rossell, S., Stuart, G. W., Shea, T. L., Innes-Brown, H., Henshall, K., ... & Egan, G. F. (2010). Reduced connectivity of the auditory cortex in patients with auditory hallucinations: a resting state functional magnetic resonance imaging study. *Psychological medicine*, 40(7), 1149-1158.

[110] Razlighi, Q. R., Habeck, C., Steffener, J., Gazes, Y., Zahodne, L. B., MacKay-Brandt, A., & Stern, Y. (2014). Unilateral disruptions in the default network with aging in native space. *Brain and behavior*, 4(2), 143-157.

[111] Franzmeier, N., Brendel, M., Beyer, L., Slemann, L., Kovacs, G. G., Arzberger, T., ... & Ewers, M. (2022). Tau deposition patterns are associated with functional connectivity in primary tauopathies. *Nature communications*, 13(1), 1362.

[112] American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.).

[113] Konradi, C., Yang, C. K., Zimmerman, E., Lohmann, K. M., Gresch, P., Pantazopoulos, H., Berretta, S., & Heckers, S. (2011). Hippocampal interneurons are abnormal in schizophrenia. *Schizophrenia research*, 131, 165 - 173.

[114] Heckers, S., & Konradi, C. (2014). GABAergic mechanisms of hippocampal hyperactivity in schizophrenia. *Schizophrenia research*, 167, 4 - 11.

[115] Videbech, P., Ravnkilde, B., Pedersen, A., Egander, A., Landbo, B., Rasmussen, N., Andersen, F., Stødkilde-Jørgensen, H., Gjedde, A., & Rosenberg, R. (2001). The Danish PET/depression project: PET findings in patients with major depression. *Psychological Medicine*, 31, 1147 - 1158.

[116] Johnston, B., Tolomeo, S., Gradin, V., Christmas, D., Matthews, K., & Steele, J. (2015). Failure of hippocampal deactivation during loss events in treatment-resistant depression. *Brain : a journal of neurology*, 138 Pt 9, 2766-76.

[117] Ghasemi, M., Navidhamidi, M., Rezaei, F., Azizikia, A., & Mehranfard, N. (2021). Anxiety and hippocampal neuronal activity: Relationship and potential mechanisms. *Cognitive, Affective, & Behavioral Neuroscience*, 22, 431 - 449.

[118] Patel, R., Spreng, R. N., Shin, L. M., & Girard, T. A. (2012). Neurocircuitry models of posttraumatic stress disorder and beyond: A meta-analysis of functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, 36(9), 2130-2142.

[119] Stramaccia, D. F., Meyer, A.-K., Rischer, K. M., Fawcett, J. M., & Benoit, R. G. (2021). Memory suppression and its deficiency in psychological disorders: A focused meta-analysis. *Journal of Experimental Psychology: General*, 150(5), 828-850.

[120] Spreng, R., & Grady, C. L. (2010). Patterns of Brain Activity Supporting Autobiographical Memory, Prospection, and Theory of Mind, and Their Relationship to the Default Mode Network. *Journal of Cognitive Neuroscience*, 22, 1112-1123.

[121] Gao, X., Yang, J., Li, C., Guo, X., Wang, Y., Qu, Z., Li, W., Liu, J., Li, Z., Li, W.-J., Hu, Z., Li, J.-J., Huang, Y., Chen, J., Wen, H., Zhao, Z., Kaiser, D., Wei, T., & Yuan, B. (2024). How do we imagine a speech? A triple network model for situationally simulated inner speech. *bioRxiv*.

[122] Whitfield-Gabrieli, S. & Ford, J. M. Default mode network activity and connectivity in psychopathology. *Annual review of clinical psychology* 8, 49–76 (2012).

[123] Ingram, R. (1990). Self-focused attention in clinical disorders: review and a conceptual model. *Psychological bulletin*, 107 2, 156-76.

[124] Fang, A., Baran, B., Beatty, C., Mosley, J., Feusner, J. D., Phan, K., Wilhelm, S., & Manoach, D. (2021). Maladaptive self-focused attention and default mode network connectivity: a transdiagnostic investigation across social anxiety and body dysmorphic disorders. *Social Cognitive and Affective Neuroscience*, 17, 645 - 654.

[125] Mullinger, K., Mayhew, S. D., Bagshaw, A., Bowtell, R., & Francis, S. (2014). Evidence that the negative BOLD response is neuronal in origin: A simultaneous EEG-BOLD-CBF study in humans. *NeuroImage*, 94, 263-274.

[126] Cole, M. W., Ito, T., Schultz, D., Mill, R., Chen, R., & Cocuzza, C. (2019). Task activations produce spurious but systematic inflation of task functional connectivity estimates. *NeuroImage*, 189, 1–18.

[127] Zhang, D.-D., Xie, H., Liu, Y., & Luo, Y. (2016). Neural correlates underlying impaired memory facilitation and suppression of negative material in depression. *Scientific Reports*, 6.

[128] Yang, W., Liu, P., Zhuang, K., Wei, D., Anderson, M. C., & Qiu, J. (2020). Behavioral and neural correlates of memory suppression in subthreshold depression. *Psychiatry Research: Neuroimaging*, 297, 111030.

Acknowledgements

There are many people without whom this project wouldn't have been possible.
First of all, my CBU supervisor.
From the day he replied to my first email to the day I left Cambridge, Akul was always able to push my understanding, push my self-imposed limits and the standard of my work, all in the kindest and friendly way possible.
It's thanks to him if I was able to delve head-first into this project and its countless details and to have learned so much from my mistakes, and for this, I'm forever grateful.
I can't express how much I'm grateful to my parents: their unconditional love, enthusiasm and support have made all this academic journey possible.
My family's trust in me has driven me in reaching whatever I thought was possible, knowing that I'd have a safe place back in our treasured home. Their unconditional support has often brought me to tears during my stay in the UK.
A special thanks to my brother, who always believed in me.
I'm grateful to my grandparents, who encouraged me in my studies and taught me that while life can often be hard, a smile doesn't have to be conditional.
I'm grateful to my best friends, whom I've shared with 10 years of pure joy.
I'm grateful to my friends in Pavia, without whom my stay there would have been certainly sadder. Sharing with you both an academic endeavour and a friendship has been extremely rewarding.
I'm then grateful to my CBU friends, who welcomed me in Cambridge and made me feel at home even at my workplace. If I've never felt lonely during my stay, it's thanks to them.
A thanks goes to all my friends around Italy, I jealously keep in my heart a lesson from each of them.
Lastly, I'll be forever grateful to my girlfriend. Her love inspires me with the courage to give it my all in every endeavour, and helps me get back on track after any setbacks.
Her unconditional love throughout the years and her ability to make me laugh in any circumstance have brought me to where I am today. A special thanks to all her family as well, who supported me at every step for the past four years.