FUNCTIONAL NEUROANATOMY OF SELF-GENERATED THOUGHT:
INVESTIGATING GENERAL BRAIN RECRUITMENT, SPECIFIC NEURAL
CORRELATES, AND NEURAL ORIGINS USING FUNCTIONAL AND
MORPHOMETRIC MAGNETIC RESONANCE IMAGING AND DIFFUSION TENSOR
IMAGING

by

KIERAN C. R. FOX

B.A., McGill University, 2008
M.A., The University of British Columbia, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

August 2016

© Kieran C.R. Fox, 2016
ABSTRACT

The human brain is not limited to the here and now, but can roam freely away from the present moment. It has evolved the remarkable capacity to conjure up the past, project itself into hypothetical futures, and envision entirely imaginary worlds and experiences. All of these capacities can be understood as variations on the theme of the brain’s ability to self-generate thoughts and imagery without any direct relation or relevance to the surrounding environment and its continuous sensory inputs. Here we examine the neural basis of such self-generated thought from a variety of angles, aiming to address numerous fundamental questions, including: general brain recruitment associated with self-generated thought; the brain basis of self-generated thought throughout the sleep cycle; specific neural correlates associated with specific types of thought; whether experimental contexts (such as the MRI scanner) influence thought content; whether specific individual patterns or ‘styles’ of thinking are associated with individual neuroanatomical heterogeneities in grey and white matter; which brain regions are necessary and sufficient for the experience of self-generated thought in wakefulness and sleep; and finally, the neuroanatomical origin sites that give rise to self-generated mental content. Substantial and meaningful (if preliminary) answers are provided to each of these questions based on six studies involving a variety of meta-analytic and empirical methods, and a combination of behavioral, functional neuroimaging, and morphometric neuroimaging data.
PREFACE

I prepared the content of this dissertation. The first several sections rely heavily on previously published material on which I was the lead (i.e., first) author and main contributor to the work. Specifically, the introductory section (Chapter 1) on self-generated thought content relies heavily on a published review of thought content in waking and dreaming (Fox et al., 2013, *Front. Hum. Neurosci.*, 7(412), 1-18), as well as further work detailing the affective content of self-generated thought (Fox et al., 2014, *Front. Psychol.*, 5(1427), 1-4). Chapter 2, a meta-analytic assessment of general brain recruitment associated with self-generated thought, relies heavily on a previously published quantitative meta-analysis of functional neuroimaging studies (Fox et al., 2015, *NeuroImage*, 111, 611-621). Chapter 3 builds on a previously published meta-analytic assessment (Fox et al., 2013, *Front. Hum. Neurosci.*, 7(412), 1-18) and review (Domhoff and Fox, 2015, *Conscious. Cogn.*, 33, 342-353) of brain recruitment during rapid eye movement (REM) sleep, as well as unpublished work in progress on the neural correlates of the other sleep stages. Chapters 4, 5, and 6 are based on newly-acquired and unpublished data collected by me and on which projects I was the lead researcher. Chapters 7 and 8 rely heavily on an under revision paper synthesizing lesion and single-neuron recording studies of self-generated thought (Fox et al., under revision).
TABLE OF CONTENTS

ABSTRACT .............................................................................................................................................. ii
PREFACE ................................................................................................................................................ iii
TABLE OF CONTENTS ........................................................................................................................ iv
LIST OF TABLES ...................................................................................................................................... viii
LIST OF FIGURES .................................................................................................................................... ix
ACKNOWLEDGMENTS .......................................................................................................................... x

CHAPTER 1 – HISTORY AND OVERVIEW OF SELF-GENERATED THOUGHT ......................... 1
  Why study the self-generated thought of the ‘resting’ brain? ......................................................... 5
    The ubiquity of self-generated thought ......................................................................................... 6
    Self-knowledge, self-improvement, and mental health ................................................................. 6
    Creativity, insight, and self-generated thought .............................................................................. 8
    Pure scientific inquiry and self-generated thought ...................................................................... 9
  The content of self-generated thought .......................................................................................... 10
  The valence of self-generated thought ......................................................................................... 13
  Cognitive control and metacognition in self-generated thought .................................................. 15

CHAPTER 2 – STUDY 1: GENERAL NEURAL CORRELATES OF SELF-GENERATED THOUGHT: A META-ANALYTIC ASSESSMENT OF FUNCTIONAL NEUROIMAGING INVESTIGATIONS .................................................................................. 16
  Introduction ........................................................................................................................................ 16
  Methods ............................................................................................................................................. 18
    Search strategy ............................................................................................................................ 18
    Study inclusion and exclusion criteria ....................................................................................... 19
    Adjustment of peak coordinates ............................................................................................... 22
    Activation likelihood estimation (ALE) meta-analysis ............................................................... 22
  Results ............................................................................................................................................. 25
  Discussion ....................................................................................................................................... 28
    What constitutes the default mode network? .............................................................................. 28
    The default mode network and self-generated thought ........................................................... 29
    Frontoparietal control network involvement in self-generated thought .................................... 30
    Coupling of the default mode and frontoparietal control networks ........................................ 31
    Other cortical areas recruited by self-generated thought ....................................................... 32
    Subcortical contributions to self-generated thought? ............................................................... 34
    Implications for clinical disorders of self-generated thought ................................................ 35
    Limitations of the meta-analysis ............................................................................................... 37
    Conclusions and future directions ............................................................................................. 38

CHAPTER 3 – STUDY 2: NEURAL CORRELATES OF SELF-GENERATED THOUGHT IN SLEEP AND DREAMING: A PRELIMINARY META-ANALYTIC ASSESSMENT OF FUNCTIONAL NEUROIMAGING INVESTIGATIONS .................................................................................. 39
  Introduction: the multiplicity of sleep .......................................................................................... 39
    Sleep mentation as self-generated thought .............................................................................. 41
    Sleep can be meaningfully dissociated into stages ................................................................. 42
    Phenomenological dissociation of sleep stages ......................................................................... 44
    Electrophysiological dissociation of sleep stages ..................................................................... 45
CHAPTER 4 – STUDY 3: SPECIFIC NEURAL RECRUITMENT ASSOCIATED WITH
DIFFERENT TYPES OF THOUGHT................................................................. 73

Introduction......................................................................................... 73
Investigating abstract qualities of thought ........................................... 73
Previous efforts to understand specific neural recruitment associated with different types of thought.................................................................................................................. 76

Methods ............................................................................................. 77
Participants .......................................................................................... 77
Experimental design ............................................................................ 77
Procedure and fMRI thought sampling paradigm ............................... 78
fMRI data acquisition ......................................................................... 80
fMRI data preprocessing and analysis .................................................. 82

Results ............................................................................................... 85
Behavioral results ............................................................................... 85
SIT: stimulus-independent thought ...................................................... 85
Orthogonality of thought qualities ....................................................... 85
Neural correlates of spontaneity and directedness in stimulus-independent thought .................................................. 86
Neural correlates of goal-relatedness in stimulus-independent thought .............................................................................. 91
Neural correlates of affective valence in stimulus-independent thought .............................................................................. 93

Discussion .......................................................................................... 95
Neural correlates of spontaneously unfolding and intentionally directed thoughts .................................................. 95
Neural correlates of goal-related stimulus-independent thoughts .............................................................................. 97
Neural correlates of positively-valenced thoughts ................................ 97
Limitations .......................................................................................... 99
Conclusions and future directions....................................................... 101

CHAPTER 5 – STUDY 4: DOES THE MRI SCANNER ENVIRONMENT INFLUENCE THOUGHT CONTENT? ................................................................. 103

Introduction ........................................................................................ 103

Methods ............................................................................................. 104
Participants ........................................................................................ 104
Morphometric neuroimaging results I: grey matter concentration

Grey matter concentration associated with frequency of intentionally directed stimulus-independent thought ................................................................. 124
Grey matter concentration associated with frequency of goal-related stimulus-independent thought ................................................................. 125
Grey matter concentration associated with frequency of positively-valenced stimulus-independent thought ......................................................... 127
Grey matter concentration associated with frequency of negatively-valenced stimulus-independent thought ......................................................... 129

Morphometric neuroimaging results II: fractional anisotropy ........................................... 132

Discussion ............................................................................................................. 134
Grey matter concentration and intentionally directed thought .............................................. 134
Grey matter concentration and goal-related thought .......................................................... 134
Grey matter concentration and the valence of thought ......................................................... 135
White matter integrity and self-generated thought tendencies .............................................. 136
Cellular foundations of macroscale morphological changes .............................................. 138
Limitations ............................................................................................................. 139

CHAPTER 7 – IDENTIFYING THE NECESSARY AND SUFFICIENT NEURAL BASIS OF
SELF-GENERATED THOUGHT: THE VALUE OF NEUROPSYCHOLOGICAL LESION
STUDIES .................................................................................................................. 141

Introduction ............................................................................................................. 141
Identifying necessary and sufficient neural substrates of self-generated thought .......... 142

CHAPTER 8 – NEURAL ORIGINS AND ONTOGENY OF SELF-GENERATED THOUGHT AS
REVEALED BY HUMAN INTRACRANIAL ELECTROPHYSIOLOGY .................................. 148
<table>
<thead>
<tr>
<th>Introduction</th>
<th>148</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neural origins of self-generated thought</td>
<td>149</td>
</tr>
<tr>
<td>The subsequent ontogeny of self-generated thought</td>
<td>154</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER 9 – SUMMARY AND CONCLUSIONS</th>
<th>159</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-generated thought in general is associated with widespread brain recruitment</td>
<td>159</td>
</tr>
<tr>
<td>Self-generated thought and imagery in sleep and dreaming share a neural substrate with waking thought</td>
<td>160</td>
</tr>
<tr>
<td>Distinct categories of thought can be resolved at the neural level with fMRI</td>
<td>160</td>
</tr>
<tr>
<td>The MRI scanner environment does not appreciably affect self-generated thought</td>
<td>161</td>
</tr>
<tr>
<td>Interindividual differences in thought content are associated with individual neuroanatomical heterogeneity</td>
<td>162</td>
</tr>
<tr>
<td>Lesions studies point to necessary neural substrates of self-generated thought</td>
<td>162</td>
</tr>
<tr>
<td>Human intracranial electrophysiology suggests a medial temporal lobe origin site for self-generated thought</td>
<td>163</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>REFERENCES</th>
<th>164</th>
</tr>
</thead>
</table>
LIST OF TABLES

TABLE 1. MAJOR AREAS OF INQUIRY IN THE COGNITIVE NEUROSCIENCE OF SELF-GENERATED THOUGHT

TABLE 2. FUNCTIONAL NEUROIMAGING STUDIES EXAMINING SELF-GENERATED THOUGHT

TABLE 3. META-ANALYTIC RESULTS: BRAIN AREAS CONSISTENTLY ACTIVATED BY MIND-WANDERING AND RELATED SELF-GENERATED THOUGHT PROCESSES

TABLE 4. FREQUENCY OF SUBJECTIVE REPORTS OF COGNITIVE ACTIVITY AND DREAMING ACROSS SLEEP STAGES

TABLE 5. NEUROCHEMICAL PROFILES OF THE VARIOUS STAGES OF SLEEP AS COMPARED TO WAKING

TABLE 6. BRAIN STATES INVESTIGATED, WITH NUMBER OF CONTRIBUTING STUDIES AND FOCI, AND NUMBER OF META-ANALYTIC RESULTS OBTAINED

TABLE 7. STUDIES INCLUDED IN META-ANALYSES OF SLEEP AND DREAMING

TABLE 8. ACTIVATIONS AND DEACTIVATIONS ASSOCIATED WITH NREM1 SLEEP

TABLE 9. ACTIVATIONS AND DEACTIVATIONS ASSOCIATED WITH NREM2 SLEEP

TABLE 10. ACTIVATIONS AND DEACTIVATIONS ASSOCIATED WITH NREM3/4 SLEEP (SWS)

TABLE 11. ACTIVATIONS AND DEACTIVATIONS ASSOCIATED WITH REM SLEEP

TABLE 12. MAJOR QUALITIES OF SELF-GENERATED THOUGHT UNINVESTIGATED AT THE NEURAL LEVEL

TABLE 13. BEHAVIORAL DATA FOR ALL FMRI RUNS

TABLE 14. PAIRWISE CORRELATIONS BETWEEN THOUGHT QUALITIES

TABLE 15. ACTIVATION PEAKS FOR SPONTANEOUSLY UNFOLDING THOUGHTS (> INTENTIONALLY DIRECTED THOUGHTS)

TABLE 16. ACTIVATION PEAKS FOR GOAL-RELATED THOUGHTS (> GOAL-UNRELATED THOUGHTS)

TABLE 17. ACTIVATION PEAKS FOR POSITIVELY-VALENCED THOUGHTS (> NEUTRAL + NEGATIVE THOUGHTS)

TABLE 18. PARTICIPANT FEATURES FOR CONTEXT COMPARISONS

TABLE 19. INDIVIDUAL PATTERNS OF THOUGHT CONTENT

TABLE 20. GREY MATTER CONCENTRATION PEAKS FOR HIGHER FREQUENCY OF INTENTIONALLY DIRECTED THOUGHTS

TABLE 21. GREY MATTER CONCENTRATION PEAKS FOR HIGHER FREQUENCY OF GOAL-RELATED THOUGHTS

TABLE 22. GREY MATTER CONCENTRATION PEAKS FOR HIGHER FREQUENCY OF POSITIVELY-VALENCED THOUGHTS

TABLE 23. GREY MATTER CONCENTRATION PEAKS FOR HIGHER FREQUENCY OF NEGATIVELY-VALENCED THOUGHTS

TABLE 24. CANDIDATE CELLULAR MECHANISMS PUTATIVELY UNDERLYING MACROSCALE STRUCTURAL DIFFERENCES IN GRAY AND WHITE MATTER

TABLE 25. SUMMARY OF HUMAN ELECTROPHYSIOLOGY STUDIES DEMONSTRATING ELICITATION OF MEMORIES, THOUGHTS, OR HALLUCINATORY, DREAM-LIKE EXPERIENCES
LIST OF FIGURES

FIGURE 1. AVERAGE AFFECT IN SELF-GENERATED THOUGHT.................................................................14
FIGURE 2. META-ANALYTIC FMRI BRAIN ACTIVATIONS.............................................................27
FIGURE 3. BRAIN NETWORKS AND META-ANALYTIC ACTIVATIONS..............................................28
FIGURE 4. EEG CORRELATES OF SLEEP.........................................................................................47
FIGURE 5. BRAIN ACTIVATIONS AND DEACTIVATIONS ACROSS SLEEP STAGES..........................63
FIGURE 6. MODEL OF BRAIN ACTIVITY ACROSS SLEEP STAGES..................................................69
FIGURE 7. THOUGHT PROBE QUESTIONS AND SEQUENCE............................................................81
FIGURE 8. ACTIVATIONS FOR SPONTANEOUSLY UNFOLDING THOUGHTS......................................88
FIGURE 9. ACTIVATIONS FOR INTENTIONALLY DIRECTED THOUGHTS..........................................89
FIGURE 10. ACTIVATIONS FOR GOAL-RELATED THOUGHTS..........................................................91
FIGURE 11. ACTIVATIONS FOR POSITIVELY-VALENced THOUGHTS...................................................93
FIGURE 12. AVERAGE THOUGHT PROPORTIONS..............................................................................103
FIGURE 13. RADAR PLOT OF THOUGHT PROPORTIONS............................................................104
FIGURE 14. STABILITY OF THOUGHT CONTENT PATTERNS ACROSS CONTEXTS........................105
FIGURE 15. WHITE MATTER TRACTS: REGIONS OF INTEREST.........................................................119
FIGURE 16. GREY MATTER CONCENTRATION AND INTENTIONALLY DIRECTED THOUGHT.........123
FIGURE 17. GREY MATTER CONCENTRATION AND GOAL-RELATED THOUGHT............................125
FIGURE 18. GREY MATTER CONCENTRATION AND POSITIVELY-VALENced THOUGHT..............127
FIGURE 19. GREY MATTER CONCENTRATION AND NEGATIVELY-VALENced THOUGHT............128
FIGURE 20. THOUGHT TENDENCIES AND MEAN FRACTIONAL ANISTROPY..............................130
FIGURE 21. BRAIN LESIONS INFLUENCING SELF-GENERATED THOUGHT.................................143
FIGURE 22. MEDIAL TEMPORAL LOBE ORIGIONS OF SELF-GENERATED THOUGHT....................149
ACKNOWLEDGMENTS

Many thanks, first and foremost, to my parents, my three brothers, and the enormous extended Fox family throughout British Columbia and around the world for their ongoing support, hospitality, and interest in my doctoral research. Huge credit is due to my supervisor for six years, Kalina Christoff, who blessed me with the greatest gift a mentor can give: absolute freedom to research anything I wanted. This freedom has led quite directly to the eclectic, but hopefully still coherent, body of work I have produced over the past six years. Next, much gratitude to my PhD thesis committee, Beck Todd, Alex Rauscher, and Steve Heine, for their time, energy, thoughtful criticisms, and most importantly, their willingness to actually read through this entire tome – as tiresome reading, I suspect, as any they’ve ever undertaken. Huge shout-outs are due also to my peers, in my lab and out, who have graciously played guru to my grasshopper and passed on the oral tradition of neuroimaging study design, analysis, and interpretation: Matt Dixon, Mike Jarrett, and Melissa Ellamil.

Many other collaborators near and far have forged my thinking, assessed my analyses, ripped up my writing, and rendered various other services less easily shoehorned into amusing alliterative(ish) phrases, all to my betterment – especial mention should be made, in no particular order, of Jess Andrews-Hanna, Bill Domhoff, Nathan Spreng, Peter Sedlmeier, Tore Nielsen, Evan Thompson, Michael Liftshitz, Sam Rumak, James Floman, Nick Fitz, Liza Solomonova, Peter Reiner, and Zac Irving.

Others are less collaborators than saviors: namely, the many technicians, administrators, and other staff at the UBC MRI Research Centre, who have made it about as easy as could reasonably be expected for me to peer into the structure and function of
living, oxygen-consuming, paramagnetic-hemoglobin-containing, human brains. Eternal gratitude to Trudy Harris, Linda James, Alex Mazur, Linda Chandler, and Laura Barlow.

It has recently been found that our lab group does research that is significantly more ‘cool’ than most other labs ($p < .05$). Our lab therefore seems to consistently attract the best and brightest undergraduate research assistants, who appear to be ignoring the heated debate surrounding $p$-values and significance testing. They may feel it a privilege to work with us, but in reality it is we who are lucky to have them. Special mention should be made of the enthusiasm, meticulousness, and acumen of Savannah Nijeboer, with whom I worked closely for years, as well as Mara L. Puertolas Lopez, Manesh Girn, and Cameron C. Parro. If I weren’t so patently less mature than all of these outstanding young researchers, I would say something along the lines of “They have kept me young at heart;” instead, let it suffice to say that I have greatly appreciated, and benefitted from, the experience of working alongside them.

Penultimately, let me thank the many friends, friends of friends, and total strangers who hopped in the MRI scanner in the name of science and advancing my CV, and of course to get pictures of their own brains. I couldn’t have done it without the participants. Also kudos to my roommates for putting up with me as I wrote this dissertation accompanied by incessant electronic music at totally unreasonable volumes and times of day.

Last, but by no means least, let us never forget who pays the bills. Many thanks to the Mind and Life Institute for providing me with a Francisco J. Varela Research Award, which generously funded the majority of the MRI scanning costs associated with the work presented in this dissertation; and to the Natural Sciences and Engineering Research Council (NSERC) of Canada, for providing me with a Vanier Canada Graduate Scholarship.
for the past three years of my doctorate, as well as awarding me a two-year post-doctoral fellowship which will allow me to continue this meandering mental adventure.
γνωθι σεαυτόν
CHAPTER 1 – HISTORY AND OVERVIEW OF SELF-GENERATED THOUGHT

The human brain is not limited to the here and now, but can roam freely away from the present moment. It has evolved the remarkable capacity to conjure up the past (Squire, 2004), project itself into hypothetical futures (Szpunar et al., 2014), and envision entirely imaginary worlds and experiences (Windt, 2010, Fox et al., 2013, Windt, 2015). All of these capacities can be understood as variations on the theme of stimulus-independence, i.e., the brain’s ability to self-generate thoughts and imagery without any direct relation or relevance to the surrounding environment and its continuous sensory inputs. Given that the least controversial perspective on the evolution of the brain argues that central nervous systems arose to more effectively allow organisms to navigate, manipulate, and anticipate their environments (Striedter, 2005), the ability of the brain to disengage from the surrounding world, and the frequency with which it exercises this capacity, are truly astounding: an estimated 30-50% of our waking hours are accompanied by self-generated and stimulus-independent thought in some form or another (Klinger and Cox, 1987, Kane et al., 2007, Killingsworth and Gilbert, 2010), and a full third of our lives is spent almost completely decoupled from the external world in various stages of sleep – all of which are accompanied by stimulus-independent cognition to some degree (Nielsen, 1999, Hobson et al., 2000, Fox et al., 2013, Domhoff and Fox, 2015).

Where do all these thoughts come from? It may be surprising that the seemingly straightforward answers “from the mind,” or “from the brain,” are in fact an incredibly recent, modern understanding of self-generated and creative thought. For nearly all of human history the source of our thoughts, particularly the most important and useful, was almost universally ascribed to divine, demonic, or otherwise external sources of some kind (Jaynes, 1976, Kuijsten,
2006, McMahon, 2013). Only in the past few centuries have human beings taken full responsibility for their own mental content, and finally localized thought to the central nervous system – laying the foundations for a protoscience of self-generated thought.

Despite this general acceptance that our own brains are the originators of this self-generated thought and imagery, cognitive neuroscience has, until recently, spent very little time or effort investigating this largely forgotten ‘other half’ of our psychological and neurophysiological functioning (Christoff, 2012). The huge accumulated body of work on task-driven and task-evoked brain function has recently been complemented, however, by a small but rapidly growing number of studies investigating the brain’s ‘intrinsic,’ resting functioning (Raichle, 2010). Aside from immense interest in resting brain metabolism (Raichle and Mintun, 2006) and the role of spontaneous neuronal activity in predictive coding (Clark, 2013), a burgeoning subfield has begun to focus on the self-generated and self-referential cognition and imagery that seem to ubiquitously accompany these ‘default’ modes of brain functioning (Gusnard et al., 2001, Raichle et al., 2001).

These ‘default’ patterns of mental activity have gone by many names, but throughout this work, we will generally use the broad term ‘self-generated thought’ (Andrews-Hanna et al., 2014) rather than the more specific ‘stimulus-independent thought.’ Whereas our own new research contributions (presented in Chapters 4-6) focus specifically on stimulus-independent thought, much of the relevant work in this field has not explicitly distinguished between stimulus-independence and -dependence in the thoughts generated by participants in studies investigating ‘mind-wandering’ (Killingsworth and Gilbert, 2010), ‘daydreaming’ (Singer, 1966), ‘task-unrelated thought’ (Allen et al., 2013), ‘spontaneous thought’ (Dumontheil et al., 2010, Ellamile et al., 2016), ‘undirected thought’ (Christoff, 2012), and ‘mentation’ in sleep and
dreaming (Nielsen, 2000). Although much of this thought can be safely assumed to be stimulus-independent, nonetheless stimulus-independence per se is best thought of as a technical term and a quality of thought that must be explicitly assessed during experience sampling. Moreover, self-generated streams of thought can interact dynamically with the environment, for instance in social situations or aesthetic experiences, sometimes making it difficult to determine where environmental influence ends and autonomously-generated thought begins (Dixon et al., 2014). A more vague term such as ‘self-generated thought’ therefore compensates for its ambiguity with its versatility, and allows us to reserve the term stimulus-independent thought for cases in which we are certain its use is justified.

In this introductory chapter we aim to lay the basic groundwork that will support the later, more detailed discussions of new meta-analytic and experimental findings. We first outline a number of reasons for studying the brain’s self-generated thought, followed by brief overviews of the content and affective valence of such thought, as well as the appreciable role played in these largely ‘undirected’ forms of thinking by cognitive control and metacognitive resources. A roadmap is presented in Table 1, which highlights major areas of inquiry in the cognitive neuroscience of self-generated thought, and outlines the 6 studies included in this dissertation that have attempted to provide some preliminary answers to these fascinating and complex problems. A final Chapter 9 provides a summary and general conclusions.
Table 1. Major areas of inquiry in the cognitive neuroscience of self-generated thought.

<table>
<thead>
<tr>
<th>Chapter (Experiment)</th>
<th>Aspect of self-generated thought</th>
<th>Central Questions</th>
<th>Principal Modality of Investigation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ch. 1</td>
<td>Content, Valence, Cognitive Control</td>
<td>What is the basic phenomenological content of self-generated thought, and how is it valenced? What role do cognitive control and metacognition play in these forms of thought?</td>
<td>First-person reports; questionnaires</td>
</tr>
<tr>
<td>Ch. 2 (Study 1)</td>
<td>General Brain Recruitment</td>
<td>What brain regions and networks are consistently recruited by self-generated forms of thought?</td>
<td>Functional neuroimaging (fMRI and PET)</td>
</tr>
<tr>
<td>Ch. 3 (Study 2)</td>
<td>General Brain Recruitment in Sleep States</td>
<td>What brain regions and networks are consistently recruited during the various stages of sleep, which differ in terms of frequency and intensity of self-generated thought and imagery?</td>
<td>Functional neuroimaging (fMRI and PET)</td>
</tr>
<tr>
<td>Ch. 4 (Study 3)</td>
<td>Specificity of Brain Recruitment</td>
<td>How does brain recruitment vary with the content and type of self-generated thought?</td>
<td>Functional neuroimaging (fMRI and PET)</td>
</tr>
<tr>
<td>Ch. 5 (Study 4)</td>
<td>Influence of Context</td>
<td>Does the surrounding environment markedly affect thought content? Specifically, does the MRI scanner have a significant influence on patterns of thought?</td>
<td>Behavioral vs. MRI scanner</td>
</tr>
<tr>
<td>Ch. 6 (Studies 5 and 6)</td>
<td>Neuroanatomical Correlates</td>
<td>Are individual differences in overall patterns of self-generated thought associated with differences in grey matter concentration and fractional anisotropy?</td>
<td>Morphometric neuroimaging (anatomical MRI and DTI)</td>
</tr>
<tr>
<td>Ch. 7</td>
<td>Necessary Substrate</td>
<td>What regions of the brain form the necessary (if not sufficient) neural substrate allowing for the self-generation of thoughts?</td>
<td>Neuropsychological lesion studies</td>
</tr>
<tr>
<td>Ch. 8</td>
<td>Origin and Ontogeny</td>
<td>How and where is thought self-generated in the brain? How do initial patterns of activity recruit and interact with other parts of the brain?</td>
<td>Human intracranial electrophysiology; magnetoencephalography</td>
</tr>
</tbody>
</table>
Why study the self-generated thought of the ‘resting’ brain?

There are countless reasons to study self-generated thought and how it is instantiated in the central nervous system, but given the difficulty often encountered by the researcher attempting to explain these reasons to peers, colleagues, and funding agencies – and the skepticism with which such explanations are often met – it is worth summarizing some major viewpoints explicitly. At least four major rationales can be articulated, two of them essentially utilitarian, the others of a more idealistic bent. The first reason is the simple ubiquity of self-generated thought: any cognitive activity occurring so frequently and automatically is probably worth investigating and understanding. The second rationale centers on the idea that self-improvement, as well as the effective treatment of mental health disorders involving dysfunctional patterns of thinking or feeling, requires self-knowledge. That is, optimizing or ameliorating our patterns of thought, feeling, and behavior first requires an understanding of what these patterns are and how the brain gives rise to them. The third notion emphasizes that, despite the obvious role of hard work, long study, and intelligence, essentially every truly novel and creative idea throughout history has arisen somehow spontaneously and in a manner beyond the control and understanding even of the thinker (Campbell, 1960, Simonton, 1999a, Simonton, 1999b). Ordinary self-generated thought, on this view, is the infinitely fertile ground from which occasionally spring all the most revolutionary and important ideas and creations of human civilization. Understanding the tails of this distribution of human creativity necessarily requires understanding its center as well, and the many ways in which our brains engage in ‘little c’ creativity in everyday life (Silvia et al., 2014). Finally, a more idealistic view argues simply that the intrinsic functioning of mind and brain is worthy of scientific interest and investigation in its
own right, regardless of whether or not any practical benefits follow. Each of these views is considered in more detail in the following sections.

The ubiquity of self-generated thought

A central impetus for the increased attention now being given to the brain basis of self-generated thought was the early recognition of its sheer quantity. Early researchers reported, and contemporary research continues to confirm, that self-generated thoughts occur ubiquitously in everyday life (Killingsworth and Gilbert, 2010), in laboratory settings (Stawarczyk et al., 2011a), and across cultures (Song and Wang, 2012). Moreover, self-generated thought is prevalent not just during minimally-demanding tasks, or periods of rest or boredom, but during virtually every imaginable activity, from personal grooming to love-making (Killingsworth and Gilbert, 2010). The sheer ubiquity of self-generated thought therefore suggests some importance to normal cognitive function, and recommends it as an object worthy of study.

Self-knowledge, self-improvement, and mental health

What we might call the meliorist viewpoint is epitomized by the famous words that were allegedly inscribed in the forecourt to the Temple of Apollo at ancient Delphi: γνωθι σεαυτόν (gnothi seauton, ‘know thyself’) (Pausanius, 1964). A meliorist stance would argue that self-improvement (for the healthy) or recovery (for those with mental health conditions) depends critically upon self-knowledge. Suboptimal or pathological patterns of thinking and behavior need first to be understood before they can be ameliorated and overcome. Presupposed in this view is the notion that cognitive-affective functioning in most, perhaps all, people is suboptimal and contains room for improvement (Fox et al., 2016) – an argument with which it is difficult to
disagree. This viewpoint is incredibly pervasive throughout many disparate facets of modern culture, underlying (tacitly or explicitly) many contemplative traditions emphasizing observation of the mind (Ouspensky, 1957, Anālayo, 2003, Kabat-Zinn, 2003, Gunaratana, 2011), as well as most psychiatric, psychoanalytic, and cognitive-behavioral therapeutic schools of thought (Rothbaum et al., 2000, Guntrip, 2008, Kohut, 2013). Indeed, so pervasive has this idea become that it is easy to forget that the high value we tend to place on self-knowledge and basic meta-awareness is essentially a cultural convention, and that self-reflective awareness of the workings of one’s own mind might have been much less prevalent, and prized, in the past – perhaps even altogether absent (Jaynes, 1976, Kuijsten, 2006).

Increasing recognition of the relevance of individual patterns of thinking to everyday well-being at the center of the spectrum, and to mental health disorders and spiritual development at its extreme poles, means that there is an increasing need to understand both the phenomenological content and neural substrates of self-generated forms of cognition. Understanding what ‘normal’ people think about, and how their brains generate such thoughts, is a useful (we would argue essential) starting point for understanding how self-generated thought can go awry, for instance in depressive rumination (Berman et al., 2010), involuntary re-experiencing in post-traumatic stress disorder (Ehlers et al., 2004, Berntsen and Rubin, 2008), and dementia (Irish et al., 2012). A basic grasp of the psychology and neurophysiology of self-generated thought has obvious relevance for spiritual paths as well (Brewer et al., 2011, Schooler et al., 2014, Eifring, 2015), but such a discussion is beyond the scope of the present work. The clinical relevance of a basic neuroscientific and neurobiological understanding of self-generated thought, however, is a theme that will recur throughout the ensuing discussions.
**Creativity, insight, and self-generated thought**

A second view credits self-generated and spontaneous forms of thought with virtually every important insight and creation throughout the history of the sciences and the arts. While the role of deliberate study, practice, and thought are of course acknowledged, the argument runs that the key insights and truly novel recombinations of ideas are, somewhat mysteriously, generated ‘in’ us but not ‘by’ us; we are the authors, but perhaps not the agents, behind our own inspirations and creative output (Campbell, 1960, Simonton, 1999a, Simonton, 1999b). Schopenhauer aptly described this process when he insisted that mental ‘rest’ must be cultivated, and space made in which great thoughts can arise of their own accord (Schopenhauer and Payne, 1974). He wrote, “Sound serious thoughts on worthy subjects... cannot be conjured up arbitrarily and at any time. All we can do is to keep the path clear for them... We need only keep the field open to sound ideas and they will come. Therefore whenever we have a free moment with nothing to do, we should… let our mind become tranquil, and then in it something good may easily arise” (pg. 54).

To be sure, novel ideas are both facilitated and constrained by our external environments, but nonetheless the remarkable yet totally commonplace creative capacity of the mind-brain needs to be recognized and understood. On this view the creations of artistic and scientific genius are simply the extremely successful examples of the kinds of thought all of us are generating virtually all of the time. We all generate many novel thoughts, images, and plans on a daily basis (Silvia et al., 2014), and an understanding of ‘Creativity’ requires an understanding of the brain’s basic capacity for autonomously generating novel content independent of its environment – despite the fact that much of this self-generated content is rather mundane.
*Pure scientific inquiry and self-generated thought*

What could be called the *scientific reductionist* viewpoint would argue that the mind and brain, like every other part of the universe, is inherently worth investigating and understanding, and that the best initial approach to such an understanding is isolation of the object of interest. ‘The’ scientific method is an ever-evolving and culturally-defined means of inquiry (Bunge, 1998), but nonetheless a series of principles of scientific investigation has long been articulated and embraced by a majority of researchers, whether explicitly or implicitly. The classic reductionist scientific approach has been to isolate the system of interest as much as possible and vary one parameter at a time in order to begin to grasp the functioning and qualities of a given object or organism. Despite the practical and philosophical limitations of this program, there is no doubt that it has been extremely successful in expanding human knowledge in virtually every field of inquiry (Capra and Luisi, 2014). Another rationale for studying the self-generated activity of the (mostly) isolated brain, then, is that observation of the mind in relative isolation might reveal much about its intrinsic functioning and proclivities. The physician John C. Lilly sought to apply this methodology to the most complex object in the known universe – the human central nervous system – when he invented the sensory deprivation tank to study the mind in isolation from its surrounding environment and interactions with others (Lilly, 1956, 1972, Lilly, 1977). Lilly explained: “During those times when it [the mind] is unencumbered by the necessities of interlock with… an external reality, its noninterlock structure can be studied. A given mind seen in pure culture by itself in profound physical isolation and in solitude is the raw material for our investigation” (Lilly, 1972, pg. xxii). This viewpoint, then, would argue that studying the mind-brain in a largely ‘unperturbed’ state is of inherent scientific interest merely for its own sake.
The corollary to all this is that the potential benefits and insights of pure research are impossible to anticipate and appreciate. Pure scientific inquiry into self-generated thinking and its brain basis might very well shed light on basic principles of central nervous system functioning or personality psychology, but such an outcome is neither necessary nor promised. An oft-repeated anecdote regarding the discovery of the relationship between electricity and magnetism by Michael Faraday (though perhaps apocryphal) aptly illustrates the point: After demonstrating that electric current created a magnetic field which could exert motive force on susceptible objects within the field (such as a nearby compass needle), a questioner allegedly asked Faraday to what possible use this quaint discovery could ever be put. Faraday is alleged to have responded: “Sir, of what use is a newborn baby?” Analogously, we might contend that the infancy of cognitive neuroscience as a field should caution us against any race to apply our meager understanding of intrinsic patterns of mental and neurophysiological functioning in quests to cure mental illness, explain personality, and so on. That is, we should study self-generated thought and its brain basis because we can, and because it is inherently interesting.

**The content of self-generated thought**

There are many sensory and more abstract qualities of thought that have been investigated by psychologists over the past few decades. We briefly review this content here to give a sense of the vast diversity of thought. Moreover, familiarity with recurring themes in the content of thought will help not only to motivate, but to interpret, neuroscientific investigations of self-generated thought: a detailed and ever-improving understanding of the content of self-generated thought helps mitigate the problem of attributing functions to these areas via reverse inference (Christoff and Owen, 2006, Poldrack, 2006). We now know, from dozens of content
studies over the past six decades, that self-generated thought content is largely audiovisual in nature (Klinger and Cox, 1987, Delamillieure et al., 2010, Stawarczyk et al., 2011a, Song and Wang, 2012); is accompanied by a wide variety of emotions (Killingsworth and Gilbert, 2010, Fox et al., 2014, Krans et al., 2015); often involves interoceptive signals and other thoughts about the body (Delamillieure et al., 2010, Diaz et al., 2013); and is typically structured as narratives involving imagined social interactions (Mar et al., 2012) centering on an individual’s current concerns, plans, and goals (Diaz et al., 2013, Klinger, 2013, Poerio et al., 2013, Stawarczyk et al., 2013). These content studies exploring the phenomenological experience of mind-wandering and other kinds of self-generated thought can therefore inform efforts to understand how diverse brain regions might be recruited in association with specific types of content, which we explore in later chapters.

Implausibility and bizarreness are also common features of self-generated thought in waking and also sleep. Although many thoughts involve relatively realistic simulations of plausible events in the external world, nonetheless a substantial number (~20% of reports) contain elements that are bizarre, implausible, or fanciful (defined as ‘departing substantially from physical or social reality’) (Klinger and Cox, 1987, Kroll-Mensing, 1992, Klinger, 2008). A more recent study has provided a general replication of earlier results: analyzing thousands of thoughts reported by 124 subjects, Kane and colleagues found that the average self-generated thought contained a moderate level of fantasy (a mean of 3.77 on a 7-point scale) (Kane et al., 2007).

Temporal orientation has also been investigated. Self-generated thought appears to involve roughly equal percentages of thoughts about the past and future (Fransson, 2006), though some studies suggest a ‘prospective bias’ toward future-oriented thoughts (Smallwood et al.,
2009, Andrews-Hanna et al., 2010, Stawarczyk et al., 2011a), and also a past-bias inducible by negative mood (Smallwood and O'Connor, 2011). Overall, however, it is clear that memories, particularly episodic ones, play a large role in self-generated thought. Many studies have reported a high prevalence (~20% or more of reports) of past-focused mind-wandering (Fransson, 2006, Smallwood et al., 2009, Andrews-Hanna et al., 2010, Smallwood and O'Connor, 2011). Indeed, one of the first studies to explore ‘resting state’ activity using PET noted the similarities between such activity and episodic memory recall, as well as the fact that subjective reports of ‘rest’ actually involved a large amount of past recollection and future planning (Andreasen et al., 1995).

Regardless of sensory modality or temporal orientation, the content of self-generated thought centers heavily on people’s current goals and concerns in life (Klinger and Cox, 1987, Klinger, 2008, McMillan et al., 2013). Further, the large percentage of self-generated thoughts centered around the present time ± 1 day (~40% in one recent study: Andrews-Hanna et al., 2010) supports the notion that thought centers on current concerns and experiences. Future-oriented thought is also incredibly common during MW (Smallwood et al., 2009, Andrews-Hanna et al., 2010, Stawarczyk et al., 2011a), further supporting a role for self-generated thought in future-planning and potentially problem-solving.

Mentalizing and simulated social interaction are also key components. Self-generated thought often involves imagined social interactions with others, as well as thoughts about the intentions and beliefs of other people (Klinger, 2008, Mar et al., 2012). This has led to the general notion that ‘mentalizing’ (i.e., thinking about the thoughts and minds of others) and social problem-solving via the consideration of hypothetical social situations may be key functional roles of self-generated thought (Buckner et al., 2008b, Andrews-Hanna, 2012,
Andrews-Hanna et al., 2014). Supporting this idea, numerous studies have found that brain activity underlying ‘theory of mind’ and mentalizing overlaps significantly with the same regions and networks implicated in self-generated thought (Spreng et al., 2009, Fox et al., 2015).

The valence of self-generated thought

Affect is ubiquitous during self-generated thought. One recent study, for instance, involving thousands of reports, found that the majority (69%) of self-generated thought reports involved emotion (positive emotion in 42.5% of reports, negative emotion in 26.5%), whereas only 31% of reports were reported to be emotionally neutral (Killingsworth and Gilbert, 2010). A prevalent viewpoint has been that self-generated thought is, overall, characterized by unhappiness or negative mood (Killingsworth and Gilbert, 2010, Wilson et al., 2014). Surprised by these claims, we recently reviewed all investigations of self-generated thought that examined affective valence, and found that in fact all studies to date have reported a mild positivity bias to self-generated thought (Fox et al., 2014, Fox et al., in preparation). This general positivity bias (Fig. 1) should not overshadow the fact that variability is high both between and within individuals, and that an incredible diversity of affect can be experienced in self-generated thought (Fox et al., 2014, Fox et al., in preparation).
Figure 1. Average reported affect across 9 studies investigating the valence of self-generated thought. Note that regardless of the particular scales, anchor points, terms, or questions employed, average valence (red dots) is always mildly positive (slightly above the midpoint on every scale). Reproduced from Fox et al. (in preparation).
Cognitive control and metacognition in self-generated thought

A lack of specific goals, and an unawareness that one is even daydreaming or has deviated from the task at hand, are typical of self-generated thought (Schooler et al., 2011, Fox and Christoff, 2015). But although self-generated thought tends to be less characterized by intentional direction and guidance and self-reflective awareness, this is not always the case. One’s impression of the ‘controllability’ of a segment of self-generated varies widely, from a sense of being able to end it at any time, to being completely absorbed and swept along (Klinger, 1978, Klinger and Cox, 1987, Kroll-Mensing, 1992, Klinger and Kroll-Mensing, 1995, Klinger, 2008). Waking self-generated thought also features a high proportion of goal-related thoughts (~50% of all thoughts; see Chapters 4 to 6), a high degree of future planning (Baird et al., 2011), a moderate (on average) amount of intentional direction of thought flow and content (~25% or more of thoughts; see Chapters 4 to 6), and relatively high levels of meta-awareness (~44% of thoughts; Christoff et al., 2009). Waking self-generated thought in general therefore involves an appreciable amount of executive and metacognitive functioning, consistent with recruitment of canonical control network regions (see Chapter 2). Collectively, these results suggest that cognitive control and metacognitive awareness in waking self-generated thought lie somewhere between the relative lucidity and self-reflectiveness of externally-directed thought and behavior, and the near-total lack of control and metacognitive nescience characteristic of cognition in sleep and dreaming (see Chapter 3). A major aim of subsequent chapters (especially Chapters 4 and 6) is to elucidate neural correlates and mechanisms that support such executive and metacognitive operations in self-generated thought.
CHAPTER 2 – STUDY 1: GENERAL NEURAL CORRELATES OF SELF-GENERATED THOUGHT: A META-ANALYTIC ASSESSMENT OF FUNCTIONAL NEUROIMAGING INVESTIGATIONS

Introduction

In parallel with the burgeoning investigation of self-generated thought, a large and influential literature has accumulated exploring the so-called ‘default mode’ network (DMN). The DMN consists of a number of functionally and anatomically interconnected brain regions that show greater activation at rest than during a variety of externally-oriented attention tasks (Shulman et al., 1997, Buckner et al., 2008a). This pattern of activity is thought to reflect a “default state of brain activity” during which “information broadly arising in the external and internal milieu is gathered and evaluated” (Raichle et al., 2001).

Given the ample evidence for the occurrence of self-generated thought during ‘rest’ states as well as minimally-demanding tasks (Singer, 1966), a reasonable supposition emerged early on in these investigations that there must be a close connection between DMN regions and mind-wandering (Andreasen et al., 1995, Raichle et al., 2001). Neuroscientific investigations into the brain basis of self-generated thought have since provided substantial empirical support for this link (e.g., (Mason et al., 2007, Christoff et al., 2009, Andrews-Hanna et al., 2010)). The intuitive appeal of this connection, together with the accumulating empirical evidence in its favor, have led cognitive neuroscientists to focus increasingly on the involvement of the DMN in self-generated forms of thought, while at the same time giving relatively little attention to the involvement of non-DMN brain areas.

Although many researchers reject a simple equivalence between self-generated thought and DMN activity, in practice the real breadth of brain recruitment during self-generated thought
has often been ignored, and a DMN-centric view often shapes the experimental design, analysis, and interpretation of neuroimaging results. Neuroimaging studies of self-generated thought often restrict analyses to DMN regions-of-interest, effectively excluding consideration of other regions. Similarly, neuroimaging studies of clinical disorders marked by dysfunctions of self-generated thought often focus their investigations on changes in DMN functioning, giving little attention to non-DMN areas (Sheline et al., 2009, Berman et al., 2010, Hamilton et al., 2011), and theoretical accounts have adopted a similar stance (e.g., (Anticevic et al., 2012, Whitfield-Gabrieli and Ford, 2012)). This prevalent experimental and theoretical bias severely limits the neuroscientific investigation of the true breadth of self-generated thought and its related clinical conditions.

Although the hypothesized connection (Gusnard et al., 2001, Raichle et al., 2001) between self-generated thought processes and activation of the DMN has now been widely validated in cognitive and clinical neuroscience (Andrews-Hanna et al., 2014, Dixon et al., 2014, Fox and Christoff, 2014, Stawarczyk and D'Argembeau, 2015), this relationship has come to overshadow the possible contribution of non-DMN regions to self-generated thought. Empirical neuroimaging work from our own group and our collaborators (Christoff et al., 2004, Christoff et al., 2009, Andrews-Hanna et al., 2010, Christoff, 2012, Ellamil et al., 2016), as well as many other independent research groups (McGuire et al., 1996, Binder et al., 1999, D'Argembeau et al., 2005, McKiernan et al., 2006, Spiers and Maguire, 2006a, Mason et al., 2007, Wang et al., 2009, Dumontheil et al., 2010, Stawarczyk et al., 2011b), has consistently shown recruitment of non-DMN regions during mind-wandering and related forms of self-generated thought. Yet the robustness of these findings has remained largely unacknowledged, and unexamined by systematic whole-brain meta-analysis. As a consequence, activity in a few selected DMN
regions-of-interest (ROIs) continues to be used as the standard, and by implication, sufficient neural indicator of self-generated thought in both healthy (Stawarczyk et al., 2011b, Vanhaudenhuyse et al., 2011) and clinical (Berman et al., 2010, Anticevic et al., 2012, Whitfield-Gabrieli and Ford, 2012) populations.

Restricting analyses to DMN regions is also problematic because the subjective variety of self-generated thought modes (which can include memory recall, planning for the future, hypothetical imaginings, mentalizing, and so on) suggest correspondingly complex and variegated neural correlates (Andrews-Hanna et al., 2013, Fox et al., 2013, Dixon et al., 2014, Fox and Christoff, 2014); a simple DMN–self-generated thought mapping is unlikely to fulfill this criterion.

Here we present a meta-analysis that specifically aimed to assess the consistency of recruitment of various regions throughout the brain during spontaneous thought processes. Rather than focusing on DMN contributions, we took a whole-brain approach to carry out a quantitative meta-analysis of the 24 functional neuroimaging studies of mind-wandering and related self-generated thought processes published to date (see Table 2). The goal is to give due consideration to the involvement of all relevant brain areas and networks.

Methods

Search strategy

To identify neuroimaging studies related to mind wandering and self-generated thought, we conducted a comprehensive and systematic search of the literature using MEDLINE (http://www.pubmed.com), Google Scholar (http://scholar.google.com), and PsycINFO (http://www.apa.org/pub/databases/psycinfo/index.aspx) for papers containing the words ‘mind-
wandering’; ‘mind wandering’; ‘spontaneous thought’; ‘stimulus-independent thought’; ‘task-unrelated thought’; or ‘daydreaming.’ From October 2014 to 1996, the year the first explicit functional neuroimaging study of spontaneous thought was published (McGuire et al., 1996). This list of articles was further refined by searching within results for studies that contained any of the words or phrases ‘magnetic resonance imaging’, ‘MRI,’ ‘positron emission tomography,’ ‘neuroimaging,’ or ‘brain’ within the title or abstract. Of the candidate studies, every abstract was read to confirm whether functional neuroimaging methods to study mind-wandering or related spontaneous thought processes was employed (see details of inclusion and exclusion criteria in the next section). The reference lists of each included study, as well as those of several major reviews (Andrews-Hanna, 2012, Christoff, 2012, Andrews-Hanna et al., 2014), were also searched, to ensure completeness.

**Study inclusion and exclusion criteria**

All studies using functional neuroimaging (i.e., functional magnetic resonance imaging (fMRI) or positron emission tomography (PET) were considered, but studies using purely morphometric neuroimaging methods to study mind-wandering (e.g., cortical thickness in (Bernhardt et al., 2014)) were not included. Using the search strategy detailed above, a total of 24 studies were collected (Table 2).

Only reports published in peer-reviewed journals were included (results from abstracts, presented talks, etc., were excluded). To be included, studies had to: (i) report specific peak foci of activation in either Talairach or Montreal Neurological Institute (MNI) space; (ii) report group results (case studies of single subjects were excluded); and (iii) involve healthy, non-clinical, ‘normal’ populations.
A fourth criterion was employed to ensure that some form of spontaneous thought was taking place: included studies had to (iv) employ some form of retrospective, online, or questionnaire self-report measure that indicated frequency or depth of spontaneous thought processes (a single exception to this criterion was allowed for an early fMRI study, because it was otherwise specifically designed to detect activations associated with spontaneous thought (Christoff et al., 2004); exclusion of this study did not appreciably alter the meta-analytic results, but its inclusion increased the power of our analysis). This criterion avoids the widespread reverse inference that is particularly present in this field (Christoff and Owen, 2006, Poldrack, 2006), where DMN activity is frequently interpreted as indirect evidence of mind-wandering-like processes (in the absence of any reports to that effect from participants themselves). This criterion also allowed for comparison of spontaneous thought frequency across multiple conditions, and therefore for a common feature across all the contrasts and studies included here (despite difference in comparison/control conditions; see below). All contrasts included in this meta-analysis represent a comparison between some state of higher spontaneous thought frequency with a corresponding period or task involving lower spontaneous thought frequency, or, similarly, regions where higher activity predicted higher rates of mind-wandering.

Furthermore, all these spontaneous thought frequency ratings are derived directly from questionnaires, experience sampling probes, and so on, rather than based on assumptions about when spontaneous thought frequency should be higher or lower. The mind-wandering measures employed included (a) post-scan retrospective reports (e.g., (McGuire et al., 1996, D'Argembeau et al., 2005); (b) online, trial-by-trial first-person reports of whether spontaneous thoughts were occurring or not (e.g., (Christoff et al., 2009, Stawarczyk et al., 2011b); and (c) questionnaires purporting to measure ‘trait’ levels of fantasy and spontaneous thought (e.g., (Wang et al., 2009).
A final method was (d) *inferential*, in that mind-wandering scores on a given block or task for a group tested only behaviorally were assumed to hold for a separate group of participants tested in the scanner (e.g., (Mason et al., 2007)). Details for each study are provided in Table 2.

As a fifth criterion studies had to (v) report whole-brain results. Studies that reported activation foci in ways other than 3D coordinate space were excluded (e.g. ROI analysis) as they are not compatible with the activation likelihood estimation (ALE) analysis approach. Further, ROI findings tend to be biased towards the DMN, reinforcing the false notion that the DMN is specifically (or exclusively) involved in mind-wandering.

Studies satisfying all five criteria (and therefore included in the meta-analysis) are listed in bold font in Table 2. A total of 10 of the 24 studies were included; failure to satisfy criterion (v) accounted for the majority of exclusions (details in Table 2). Two studies (Hasenkamp et al., 2012, Kucyi et al., 2013) met all five criteria, but involved other potentially confounding factors. The first (Hasenkamp et al., 2012) employed a population of expert meditation practitioners, who may have altered frequency of mind-wandering (Mrazek et al., 2013), as well as altered brain activity during periods of rest with high levels of mind-wandering (Brewer et al., 2011). The second study (Kucyi et al., 2013) examined the relationship between mind-wandering and pain, administering painful electrical nerve stimulation during the sessions where mind-wandering was reported, again presenting a potential confound. To avoid these confounds affecting our meta-analytic results these studies were excluded (although their inclusion did not appreciably impact the present findings).

One study by Spiers and Maguire (2006a) was followed up with a similar study using the same subjects, dataset, and baseline condition, but somewhat different spontaneous mental activity as the condition of interest (Spiers and Maguire, 2006b). As inclusion of both studies
would involve completely overlapping, non-independent samples and data, only one study was included. The former study involved spontaneous mentalizing, which is very common during mind-wandering and related forms of spontaneous cognition (Fox et al., 2013), and so was retained. The latter study instead investigated spontaneous route-planning during virtual navigation, which bears much less resemblance to the various other forms of spontaneous thought examined here, and so was not included in the meta-analysis.

*Adjustment of peak coordinates*

Some included studies reported results in Talairach coordinates. For consistency, we converted all coordinates to MNI space using non-linear transformations in the WFU PickAtlas software package (Maldjian et al., 2003). A small but substantial number of peak foci \( n = 16 \) were reported at coordinates that lay outside of the template brain images used within the activation likelihood estimation meta-analysis software, GingerALE. To avoid data loss, we performed a minimum linear translation of these peak coordinates in MNI space to fit within the GingerALE template mask. In all cases these translations were minimal (mean = 3.69 mm; range: 1–10 mm), and in no case did adjusted peak foci subsequently fall within a different brain region from that listed in the results tables of the original publication. Full details of these translations, including the original and adjusted foci, are available in the online supplementary materials accompanying the original publication (Fox et al., 2015).

*Activation likelihood estimation (ALE) meta-analysis*

We used a quantitative, random-effects meta-analytic method known as activation likelihood estimation (ALE) (Turkeltaub et al., 2002, Laird et al., 2005, Eickhoff et al., 2009,
Eickhoff et al., 2012, Turkeltaub et al., 2012) implemented in the software program GingerALE 2.3.1 (San Antonio, TX: UT Health Science Center Research Imaging Institute). The most recent ALE algorithm tests for above-chance clustering of peak foci from different experiments included in the meta-analysis (Eickhoff et al., 2009, Eickhoff et al., 2012) by comparing actual activation foci locations/clustering with a null distribution that includes the same number of peak foci distributed randomly throughout the brain’s grey matter. Included activation foci were smoothed using a full-width half maximum (FWHM) Gaussian kernel dependent on the sample size (subjects) of the experiment from which foci were drawn (larger sample -> smaller smoothing kernel – empirically determined by (Eickhoff et al., 2009, Eickhoff et al., 2012)). Resulting statistical maps show clusters where convergence between activation foci is greater than would be expected by chance (i.e., if foci from each experiment were distributed independently).

We analyzed a total of 157 foci from 10 neuroimaging studies (studies in bold font in Table 2). Statistical maps were thresholded using a false discovery rate (FDR – see (Genovese et al., 2002)) of \( q = .05 \) and a cluster threshold of \( k = 1000 \text{ mm}^3 \). We used this relatively liberal statistical threshold, but high cluster threshold, in order to better visualize results and to ensure that cross-study convergence was not neglected in this relatively small-sample meta-analysis. Similar results were found at more conservative statistical thresholds. Region classifications follow those indicated in the Multi-Image Analysis GUI (‘Mango’) image-viewing software (UT Health Science Center Research Imaging Institute) used to visualize the meta-analytic results. The Duvernoy neuroanatomical atlas was also consulted (Duvernoy et al., 1991).
Table 2. Functional neuroimaging studies examining self-generated thought.

<table>
<thead>
<tr>
<th>Study</th>
<th>Analysis</th>
<th>Peak foci</th>
<th>Design</th>
<th>N</th>
<th>Mind-wandering measure</th>
<th>Thought type</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGuire et al., 1996</td>
<td>WB &amp; ROI</td>
<td>5</td>
<td>Blocked</td>
<td>5/6</td>
<td>Retrospective</td>
<td>Verbal stimulus-independent thoughts during rest</td>
</tr>
<tr>
<td>Binder et al., 1999</td>
<td>WB</td>
<td>8</td>
<td>Blocked</td>
<td>14</td>
<td>Inferential</td>
<td>Task-unrelated thought during rest</td>
</tr>
<tr>
<td>Christoff et al., 2004</td>
<td>WB</td>
<td>15</td>
<td>Blocked</td>
<td>12</td>
<td>Assumptive</td>
<td>Spontaneous thought during rest</td>
</tr>
<tr>
<td>D'Argembeau et al., 2005</td>
<td>WB</td>
<td>9</td>
<td>Blocked</td>
<td>13</td>
<td>Retrospective</td>
<td>Spontaneous thought during rest</td>
</tr>
<tr>
<td>McKiernan et al., 2006</td>
<td>ROI</td>
<td>–</td>
<td>Blocked</td>
<td>30</td>
<td>Inferential</td>
<td>Task-unrelated thought during auditory task</td>
</tr>
<tr>
<td>Spiers &amp; Maguire, 2006</td>
<td>WB &amp; ROI</td>
<td>24</td>
<td>ER</td>
<td>20</td>
<td>Retrospective (Online)</td>
<td>Spontaneous mentalizing during navigation task</td>
</tr>
<tr>
<td>Mason et al., 2007</td>
<td>WB</td>
<td>20</td>
<td>Blocked</td>
<td>19</td>
<td>Inferential and Questionnaire</td>
<td>Mind-wandering during highly practiced task</td>
</tr>
<tr>
<td>Christoff et al., 2009</td>
<td>WB</td>
<td>34</td>
<td>ER</td>
<td>15</td>
<td>Online</td>
<td>Mind-wandering during SART</td>
</tr>
<tr>
<td>Wang et al., 2009</td>
<td>WB &amp; ROI</td>
<td>8</td>
<td>Blocked</td>
<td>13</td>
<td>Questionnaire</td>
<td>Spontaneous thought during rest</td>
</tr>
<tr>
<td>Andrews-Hanna et al., 2010</td>
<td>ROI; seed-based FC</td>
<td>–</td>
<td>Blocked</td>
<td>30/139</td>
<td>Surprise retrospective</td>
<td>Spontaneous thought during passive fixation</td>
</tr>
<tr>
<td>Dumontheil et al., 2010</td>
<td>WB</td>
<td>20</td>
<td>Blocked</td>
<td>16</td>
<td>Retrospective</td>
<td>Task-unrelated thought during various simple tasks</td>
</tr>
<tr>
<td>Stawarczyk et al., 2011</td>
<td>ROI/Supp. WB</td>
<td>14</td>
<td>ER</td>
<td>22</td>
<td>Online</td>
<td>Mind-wandering during SART</td>
</tr>
<tr>
<td>Vanhaudenhuyse et al., 2011</td>
<td>ROI</td>
<td>–</td>
<td>ER</td>
<td>22</td>
<td>Online</td>
<td>Intensity of internal awareness during rest</td>
</tr>
<tr>
<td>Christoff, 2012</td>
<td>ROI; seed-based FC</td>
<td>–</td>
<td>ER</td>
<td>15</td>
<td>Online</td>
<td>Mind-wandering during SART</td>
</tr>
<tr>
<td>Hasenkamp et al., 2012</td>
<td>WB</td>
<td>–</td>
<td>ER</td>
<td>14</td>
<td>Online</td>
<td>Mind-wandering during meditation</td>
</tr>
<tr>
<td>Hasenkamp &amp; Barsalou, 2012</td>
<td>ROI; seed-based FC</td>
<td>–</td>
<td>ER</td>
<td>14</td>
<td>Online</td>
<td>Mind-wandering during meditation</td>
</tr>
<tr>
<td>Allen et al., 2013</td>
<td>ROI</td>
<td>–</td>
<td>ER</td>
<td>21/21</td>
<td>Online</td>
<td>Task-unrelated thought during EAT</td>
</tr>
<tr>
<td>Kucyi et al., 2013</td>
<td>WB</td>
<td>–</td>
<td>Blocked</td>
<td>51</td>
<td>Online</td>
<td>Mind-wandering during painful stimuli</td>
</tr>
<tr>
<td>Moss et al., 2013</td>
<td>Functional localizer-defined ROIs</td>
<td>–</td>
<td>Blocked</td>
<td>15</td>
<td>Retrospective</td>
<td>Mind-wandering during reading</td>
</tr>
<tr>
<td>Smallwood, Tipper, et al., 2013</td>
<td>WB</td>
<td>–</td>
<td>ER</td>
<td>16</td>
<td>Assumptive</td>
<td>Stimulus-independent thought during various tasks</td>
</tr>
<tr>
<td>Smallwood, Gorgolewski, et al., 2013</td>
<td>ROI</td>
<td>–</td>
<td>Blocked</td>
<td>42</td>
<td>Online</td>
<td>Mind-wandering during reading</td>
</tr>
<tr>
<td>Gorgolewski et al., 2014</td>
<td>WB</td>
<td>–</td>
<td>Blocked</td>
<td>166</td>
<td>Retrospective</td>
<td>Self-generated thought during rest</td>
</tr>
<tr>
<td>Kucyi &amp; Davis, 2014</td>
<td>ROI</td>
<td>–</td>
<td>Blocked</td>
<td>51</td>
<td>Online</td>
<td>Mind-wandering during painful stimuli</td>
</tr>
<tr>
<td>Tusche et al., 2014</td>
<td>ROI</td>
<td>–</td>
<td>ER</td>
<td>30</td>
<td>Online</td>
<td>Self-generated thought during rest</td>
</tr>
</tbody>
</table>

All studies were conducted using fMRI, except McGuire et al. (1996) and D’Argembeau et al. (2005), which used PET. N = sample size (experimental/control group, if applicable). Studies employing whole-brain analyses that were
included in the meta-analysis are listed in **bold** font. The number of peak activation foci contributed from each study is indicated in the third column. EAT: Error Awareness Task; ER: event-related; FC: functional connectivity; ROI: region of interest; SART: Sustained Attention to Response Task; Supp.: supplemental; WB: whole-brain. 

See Methods for a discussion of mind-wandering measures.

**Results**

We found 13 regions to be reliably involved in mind wandering and self-generated thought (Table 3; Fig. 2). These regions included all major hubs of the DMN: rostromedial prefrontal cortex, medial prefrontal cortex/anterior cingulate cortex, precuneus/posterior cingulate cortex, bilateral inferior parietal lobule, and left medial temporal lobe/parahippocampal cortex (extending somewhat into the cerebellum). We also found significant clusters outside the DMN, including dorsal anterior cingulate cortex, right rostrolateral prefrontal cortex, left ventrolateral prefrontal cortex, secondary somatosensory cortex, left temporopolar cortex, left mid insula, and the left lingual gyrus.
Table 3. Meta-analytic results: brain areas consistently activated by mind-wandering and related self-generated thought processes.

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster Size (mm$^3$)</th>
<th>Peak ALE Value</th>
<th>Peak in MNI Space (x, y, z)</th>
<th>Brodmann Area</th>
<th>Studies Contributing to Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Frontal lobe</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dorsal anterior cingulate cortex</td>
<td>2920</td>
<td>0.0144</td>
<td>-5, 27, 39 (Fig. 2a)</td>
<td>32</td>
<td>a, e, g, h</td>
</tr>
<tr>
<td>R lateral prefrontal cortex</td>
<td>1656</td>
<td>0.0085</td>
<td>45, 43, -8 (Fig. 2e,g)</td>
<td>10/46</td>
<td>d, i</td>
</tr>
<tr>
<td>Rostromedial prefrontal cortex</td>
<td>1256</td>
<td>0.0081</td>
<td>3, 61, 13 (Fig. 2b)</td>
<td>10/9</td>
<td>e, f</td>
</tr>
<tr>
<td>Medial prefrontal cortex; anterior cingulate cortex</td>
<td>1080</td>
<td>0.0089</td>
<td>4, 42, 3 (Fig. 2b)</td>
<td>24/32</td>
<td>f, g</td>
</tr>
<tr>
<td>L ventrolateral prefrontal cortex</td>
<td>1032</td>
<td>0.0126</td>
<td>-42, 29, -12 (Fig. 2d)</td>
<td>47/11</td>
<td>e, j</td>
</tr>
<tr>
<td><strong>Parietal Lobe</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precuneus; posterior cingulate cortex</td>
<td>3784</td>
<td>0.0123</td>
<td>-8, -56, 39 (Fig. 2a)</td>
<td>7/31</td>
<td>b, d, f, g, h, j</td>
</tr>
<tr>
<td>L inferior parietal lobule; angular gyrus</td>
<td>2616</td>
<td>0.0092</td>
<td>-46, -72, 25 (Fig. 2d)</td>
<td>39</td>
<td>b, e, f, g, h</td>
</tr>
<tr>
<td>R inferior parietal lobule; supramarginal gyrus</td>
<td>1808</td>
<td>0.0098</td>
<td>56, -51, 33 (Fig. 2f)</td>
<td>40/39</td>
<td>d, f, i</td>
</tr>
<tr>
<td>R secondary somatosensory cortex</td>
<td>1416</td>
<td>0.0097</td>
<td>24, -39, 56 (Fig. 2i)</td>
<td>5/40</td>
<td>c, e, g</td>
</tr>
<tr>
<td><strong>Temporal Lobe</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L parahippocampus</td>
<td>3496</td>
<td>0.0196</td>
<td>-27, -37, -18 (Fig. 2h)</td>
<td>36</td>
<td>b, c, f, g, h</td>
</tr>
<tr>
<td>L mid insula/superior temporal gyrus</td>
<td>2608</td>
<td>0.0116</td>
<td>-50, -1, -5 (Fig. 2d,g)</td>
<td>13/22</td>
<td>f, g, i</td>
</tr>
<tr>
<td>L temporopolar cortex</td>
<td>1320</td>
<td>0.0090</td>
<td>-35, 10, -26 (Fig. 3)</td>
<td>38</td>
<td>c, g, j</td>
</tr>
<tr>
<td><strong>Occipital Lobe</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L lingual gyrus</td>
<td>1936</td>
<td>0.0092</td>
<td>-15, -66, 5 (Fig. 2c)</td>
<td>19/18/30</td>
<td>c, e, g, j</td>
</tr>
</tbody>
</table>

Significant meta-analytic clusters, with peak coordinates, for self-generated thought > control conditions. BA: Brodmann area. a: McGuire et al., 1996; b: Binder et al., 1999; c: Christoff et al., 2004; d: D’Argembeau et al., 2005; e: Spiers & Maguire, 2006; f: Mason et al., 2007; g: Christoff et al., 2009; h: Wang et al., 2009; i: Dumontheil et al., 2010; j: Stawarczyk et al., 2011.
Figure 2. Peaks of each meta-analytic cluster of activation associated with self-generated thought processes. Peaks for each of 13 significant meta-analytic clusters of activation associated with mind-wandering and related self-generated thought processes. (a) Posterior cingulate cortex/precuneus and dorsal anterior cingulate cortex. (b) Medial prefrontal cortex/anterior cingulate cortex and rostromedial prefrontal cortex. (c) Left lingual gyrus (largest cluster). (d) From left to right: left inferior parietal lobule/angular gyrus; left mid-insula; and left ventrolateral prefrontal cortex. (e) Right rostrolateral prefrontal cortex. (f) Right inferior parietal lobule/supramarginal gyrus. (g) Left temporopolar cortex and mid-insula; right rostrolateral/dorsolateral prefrontal cortex. (h) Left parahippocampus/medial temporal lobe, extending into the cerebellum. (i) Right secondary somatosensory cortex. Color bar indicates ALE likelihood values (see Methods). Stereotactic coordinates (x, y, z) are in MNI space. L: left; R: right.
Figure 3 Meta-analytic activations for self-generated thought. Cortical mapping of significant meta-analytic clusters associated with mind-wandering and related self-generated thought processes (green clusters) juxtaposed with outlines of the default mode network (blue) and the frontoparietal control network (red). Note that spontaneous thought activations overlap considerably with both networks, but also include regions beyond both networks (highlighted in Figure 1). Default mode network and frontoparietal control network masks based on Yeo et al. (2011). Reproduced from Fox et al. (2015).

Discussion

What constitutes the default mode network?

Understanding the extent to which self-generated thought recruitment overlaps with, and extends beyond, the DMN naturally requires a definition of what brain areas constitute the DMN. Because the DMN is functionally (rather than anatomically) defined, and because numerous
different methods can be used to derive this functional definition, its boundaries vary somewhat depending on the methods and contrasts employed.

Early studies using positron emission tomography compared blocked periods of ‘rest’ with blocked task periods (Andreasen et al., 1995, Shulman et al., 1997, Raichle et al., 2001). Based on these studies, the DMN came to be defined as a set of regions exhibiting decreases in activity, or deactivations, during a variety of externally-directed tasks compared to periods of rest (Shulman et al., 1997, Mazoyer et al., 2001, Raichle et al., 2001, Buckner et al., 2008b). Later work using briefer temporal epochs (event-related analyses) suggested a similar constellation of regions (Kirchoff et al., 2005), while more recent work has begun to define the DMN based on its pattern of intrinsic resting state functional connectivity (i.e., temporally correlated fluctuations in activity across different brain areas) (Greicius et al., 2003, Yeo et al., 2011).

Comparing findings across these three methods has allowed for a tentative delineation of ‘canonical’ DMN regions, which include the medial prefrontal cortex, posterior cingulate cortex, temporoparietal junction (and nearby inferior parietal lobule), and parts of lateral temporal cortex (Buckner et al., 2008b, Andrews-Hanna et al., 2014). Other regions are also sometimes included, most commonly the medial temporal lobe (especially the hippocampus) and ventrolateral prefrontal cortex (BA 47) (Buckner et al., 2008b, Andrews-Hanna et al., 2014).

The default mode network and self-generated thought

Since the beginning of investigation into DMN functions, this network has been linked to mental processes such as spontaneous self-referential processing, memory recall, and inner simulations of behavior (Andreasen et al., 1995, Gusnard et al., 2001, Raichle et al., 2001). A
growing body of work focusing specifically on self-generated thought has bolstered this connection: both retrospective (Mason et al., 2007, Andrews-Hanna et al., 2010) and online, trial-by-trial (Christoff et al., 2009) self-reported mind-wandering predict increased activity in DMN regions. Further, self-reported depth (or ‘intensity’) of internally-directed thought predicts higher activation in DMN areas (Vanhaudenhuyse et al., 2011); and spontaneous mentalizing (i.e., mind-wandering about the thoughts or intentions of others) also recruits the DMN (Spiers and Maguire, 2006a).

Regardless of the specific DMN borders used, a strong connection between the DMN and spontaneous thought emerges. Indeed, all studies of self-generated thought published so far employing an agnostic, whole-brain approach (McGuire et al., 1996, Binder et al., 1999, Christoff et al., 2004, D'Argembeau et al., 2005, Spiers and Maguire, 2006a, Mason et al., 2007, Christoff et al., 2009, Wang et al., 2009, Dumontheil et al., 2010, Stawarczyk et al., 2011b) have shown recruitment of at least some DMN regions, including medial prefrontal cortex (BA 10/32), posterior cingulate cortex/precuneus (BA 7/31), parahippocampal cortex (BA 36), and temporoparietal junction/inferior parietal lobule (BA 39/40). Our quantitative meta-analysis corroborates these findings, strongly supporting a link between the DMN and self-generated thought (Table 3; Fig. 2 and 3). Putative roles for DMN hubs and subsystems in self-generated thought are discussed at length elsewhere (Andrews-Hanna et al., 2014, Smallwood and Schooler, 2014).

**Frontoparietal control network involvement in self-generated thought**

Mind-wandering and related self-generated thought processes consistently recruited many regions of the frontoparietal control network (Vincent et al., 2008, Yeo et al., 2011), including the dorsal anterior cingulate cortex (BA 32), right rostrolateral prefrontal cortex (BA 10/46),
right anterior inferior parietal lobule (BA 39/40), and precuneus (BA 7) (Table 3; Figs. 2 and 3). The frontoparietal control network plays a central role in goal-directed cognition (Niendam et al., 2012). This role extends from classic measures of executive functioning to personally relevant goal-directed tasks, such as planning for the future (Spreng et al., 2010, Gerlach et al., 2014). In the latter capacity, it tends to be activated in unison with DMN areas, so we discuss its potential role in spontaneous thought in the following section, which addresses DMN-frontoparietal control network coupling in the generation of thought.

**Coupling of the default mode and frontoparietal control networks**

Many meta-analytic clusters fell within the borders of either the default mode (Shulman et al., 1997, Raichle et al., 2001, Buckner et al., 2008b) or fronto-parietal control network (Vincent et al., 2008, Yeo et al., 2011, Spreng et al., 2013). We visualize the extent of this overlap by plotting our meta-analytic clusters on an ‘inflated’ brain template with border outlines for both networks in Figure 3. The borders of the two networks are based on previously published, large-scale (n = 1000) resting state datasets (Yeo et al., 2011), whereas the green clusters of activation represent our whole-brain meta-analytic results (from Table 3; Fig. 2).

Although our meta-analytic results cannot directly address the temporal correlation of these brain activations, such results are consistent with observations of temporal coupling of DMN and fronto-parietal control network areas (e.g., (Spreng et al., 2010, Christoff, 2012)). Self-generated thought is not the only cognitive process, however, to exhibit co-activation of, and functional connectivity (i.e., temporal coupling) between, the default mode and frontoparietal control networks. Similar findings have been observed in autobiographical memory recall (Spreng et al., 2009), mentalizing or ‘theory of mind’ (Spiers and Maguire,
2006a, Spreng et al., 2009), creative generation and evaluation of artwork (Ellamil et al., 2012), and goal-directed prospection (Spreng et al., 2009, Gerlach et al., 2014). Intriguingly, memory recall, mentalizing, creative recombination of ideas, and thoughts about the future mark the dominant themes of self-generated thought content from the subjective point of view (reviewed in (Klinger, 2008, Fox et al., 2013)), making these overlapping brain activations intriguing avenues for future research. Resting-state functional connectivity analysis has revealed that there is an extensive intrinsic architecture to support dynamic interactions between the default and frontoparietal control networks (Spreng et al., 2013). The coupled recruitment of these two networks might represent a dynamic interplay whereby executive control regions guide, evaluate, and select among the various self-generated streams of thoughts, memories, and imaginings offered up to consciousness by the DMN (Andrews-Hanna et al., 2014, Fox and Christoff, 2014, Fox et al., 2016).

Other cortical areas recruited by self-generated thought

Areas outside the default and frontoparietal control networks also need to be given consideration. The temporopolar cortex (McGuire et al., 1996, Christoff et al., 2004, Christoff et al., 2009, Dumontheil et al., 2010), the insula (Christoff et al., 2004, Mason et al., 2007, Christoff et al., 2009, Dumontheil et al., 2010, Stawarczyk et al., 2011b), and the lingual gyrus/medial occipital cortex are all relatively consistently activated during self-generated thought (Fig. 2), but do not clearly belong to either of these two networks. Further, our meta-analysis also suggests consistent recruitment of secondary somatosensory cortices (Fox et al., 2015). What role could these regions be playing?
A detailed and ever-improving understanding of the content of self-generated thought helps mitigate the problem of attributing functions to these areas via inverse inference (Christoff and Owen, 2006, Poldrack, 2006), allowing some informed (if preliminary) hypotheses to be put forward. We now know, from dozens of content studies over the past six decades, that spontaneous thought content is largely audiovisual in nature (Klinger and Cox, 1987, Delamillieure et al., 2010, Stawarczyk et al., 2011a, Song and Wang, 2012); is accompanied by a wide variety of emotions (Killingsworth and Gilbert, 2010, Fox et al., 2014, Krans et al., 2015); often involves interoceptive signals and other thoughts about the body (Delamillieure et al., 2010, Diaz et al., 2013); and is typically structured as narratives involving imagined social interactions (Mar et al., 2012) centering on an individual’s current concerns, plans, and goals (Diaz et al., 2013, Klinger, 2013, Poerio et al., 2013, Stawarczyk et al., 2013). These content studies exploring the phenomenological experience of mind-wandering and other kinds of spontaneous thought can therefore inform efforts to understand how diverse brain regions might be recruited in association with specific types of content. For instance, the prominent recruitment of the medial occipital cortical regions, key to visual imagery and the visual component of dreaming (Solms, 1997), are a plausible substrate for the ubiquitous visual imagery in spontaneous thought. Similarly, consistent recruitment of the insula and secondary somatosensory cortex, areas key to interoception (Craig, 2009) and exteroception (Penfield and Boldrey, 1937), respectively, is in line with the prominence of body-related thought content (Delamillieure et al., 2010, Diaz et al., 2013). Temporopolar cortex recruitment could be linked to the ubiquity of conscious emotional experience and mentalizing (Olson et al., 2007), and the prevalence of positive affect in particular (Fox et al., 2014) might be expected to recruit subcortical basal ganglia structures involved in reward (see next section). Finally, the frequency
of future-oriented thought (Andrews-Hanna et al., 2010, Stawarczyk et al., 2013) involving planning and goal-related ideation would suggest the recruitment of prefrontal executive areas critical for these capacities – and this is observed in rostrolateral prefrontal cortex, dorsal anterior cingulate cortex, and possibly other lateral prefrontal areas.

These functional roles are largely speculative, of course, and we intend them merely as suggestions. Nonetheless, note that the wide breadth of brain recruitment during spontaneous thought aligns well with the wide variety of phenomenological content. The content dimensions of spontaneous thought are reviewed in more detail by Fox et al. (2013) and Klinger (2008), and further discussion of the possible functional roles of these cortical regions and their relation to content is provided by Domhoff & Fox (2015) and Fox et al. (2015).

Subcortical contributions to self-generated thought?

A corticocentric bias pervades much of cognitive neuroscience (Parvizi, 2009), and this is true too of DMN-centric views of self-generated thought. Empirical reports have shown self-generated thought to recruit numerous subcortical regions, including the cerebellum (Spiers and Maguire, 2006a, Christoff et al., 2009, Stawarczyk et al., 2011b), the putamen and claustrum (Mason et al., 2007), the caudate nucleus (Christoff et al., 2009, Stawarczyk et al., 2011b), the thalamus (Christoff et al., 2009, Stawarczyk et al., 2011b), and the hypothalamus (Stawarczyk et al., 2011b). Although no subcortical clusters emerged in our quantitative meta-analysis (Fox et al., 2015), this could have been for a number of reasons – such as low statistical power due to the small number of studies conducted to date, or because of diversity in the type of thought content investigated in each study. Continuing research may show that certain subcortical structures are consistently recruited during self-generated thought, or are recruited specifically depending on
thought content or affect. The recruitment of the basal ganglia, for instance, could be related to the highly automatized and habitual nature of many streams of spontaneous thought (Fox et al., 2016), or possibly the rewarding aspect of the many emotionally positive thoughts reported (Fox et al., 2014). The potential involvement of the cerebellum also raises intriguing possibilities: although traditionally considered mainly a motor area, more recent conceptions see it as playing a critical role in ‘higher’ cognitive processes (Middleton and Strick, 1994, Bellebaum and Daum, 2007, Stoodley and Schmahmann, 2009), especially planning and simulating future behaviors (Houk and Wise, 1995, Thach, 1996, Schmahmann and Sherman, 1998). Cerebellum recruitment might therefore occur preferentially when self-generated thought involves complex planning and simulations. These suggestions are of course speculative at this early stage; the role (if any) played by subcortical brain regions is further explored in our fMRI (Chapter 4) and morphometric MRI (Chapter 6) investigations, below.

*Implications for clinical disorders of self-generated thought*

Above, we discussed the idea that the DMN (especially medial temporal lobe structures) may contribute to the actual arising of memories and the stream of associated thoughts and imaginings. Higher executive areas may, on the other hand, subsequently evaluate, monitor, and guide the flow of these thoughts toward desirable and beneficial outcomes (Fox and Christoff, 2014, Fox et al., 2016). Given the evidence that self-generated thought involves a coupling of both DMN and frontoparietal control network regions in normal, healthy subjects, important implications arise for clinical disorders of self-generated thought.

For instance, preferential (or over-) activation of DMN regions in isolation, with minimal or attenuated contributions from executive areas, might characterize dysfunctional
forms of self-generated thought, such as depressive rumination or the spontaneous and debilitating re-experiencing characteristic of post-traumatic stress disorder (Fox et al., under review). Conversely, an over-involvement of frontoparietal control areas might characterize conditions like obsessive-compulsive disorder, where fleeting thoughts and urges are acted upon repeatedly and with determination. If the distributed neural basis of spontaneous thought is not appreciated, and DMN regions are preferentially or exclusively investigated, links between such clinical disorders of spontaneous thought and non-DMN brain regions might be missed.

Two highly-cited studies of depression (Berman et al., 2010, Hamilton et al., 2011) exemplify this potential pitfall: both found links between DMN activity and frequency of depressive rumination, but neither explored potential links with non-DMN areas, particularly executive areas of the frontoparietal control network. Yet one of these groups’ meta-analysis (Hamilton et al., 2012) has shown that several regions beyond the DMN show abnormal activity in depression, including the dorsolateral prefrontal cortex, insula, and dorsal anterior cingulate cortex (note that the two latter regions are consistently recruited by self-generated thought; see Results, above). In contrast, a recent study that did explore whole-brain relationships (Hach et al., 2014) found that depressed patients showed altered patterns of functional connectivity across numerous non-DMN areas, including dorsal anterior cingulate cortex, as well as both dorsolateral and rostrolateral prefrontal cortex (all regions of the frontoparietal control network).

Although the study investigated deliberate past- and future-oriented thought, and did not directly address ‘spontaneous’ forms of cognition, similar differences in neural recruitment could also characterize the spontaneous thoughts of depressed patients. Such possibilities are clearly worthy of further investigation with an unbiased, whole-brain approach, not only for depression but also other clinical conditions characterized by disorders of spontaneous thought.
Limitations of the meta-analysis

A number of limitations of the present meta-analytic methods, as well as of the primary research literature meta-analyzed, should be kept in mind. First, differences in baselines and other control conditions across studies are a source of heterogeneity. Collapsing across numerous different baselines or control conditions is a common, indeed inevitable, practice in any large meta-analysis of functional neuroimaging studies (Svoboda et al., 2006, Wager et al., 2007, Salimi-Khorshidi et al., 2009, Wager et al., 2009). Nonetheless, the heterogeneity attributable to the many different study designs concatenated here should be kept in mind.

Another concern is the wide variety of questionnaire and thought-sampling methods employed to determine the content and frequency of spontaneous thought. Although we would argue that subtle distinctions between different kinds of spontaneous thought are premature at this point, it should be acknowledged that various distinctive forms of spontaneous cognition might eventually be differentiable both phenomenologically, and potentially also at the neural level. The present meta-analysis, collapsing over these potentially distinguishable forms of spontaneous cognition, is meant merely as a preliminary effort at synthesis.

A third point is the fairly small sample size of studies included in the meta-analysis. Compared with processes like memory or attention, very little neuroimaging research has investigated spontaneous thought processes, and of this already small pool many studies have not employed agnostic, whole-brain analyses suitable for inclusion in a quantitative meta-analysis. As such, our meta-analytic results are based on a relatively small (if rigorous, and rigorously selected) set of studies, and should therefore be interpreted with caution. Future work may both extend and further clarify the results presented here.
Conclusions and future directions

The primary aim of this meta-analysis was to provide a quantitative overview of the brain regions and networks consistently recruited by mind-wandering and related forms of spontaneous thought. We found that a wide variety of regions throughout all lobes and both hemispheres appear to be reliably recruited. Most of these regions fall within boundaries of the default mode or frontoparietal control networks, suggesting that these networks, as presently conceptualized, form an important part of the neural signature accompanying, self-generated thought of various kinds. A number of other reliably activated regions are not so easily subsumed under either of these networks (e.g., the insula, somatosensory cortices, and the lingual gyrus), suggesting that future work may need to look beyond the typical delineation of these two networks’ contributions in order to construct a full account of the neural recruitment associated with spontaneous thought.

The study of deliberate, goal-oriented, and externally-directed mental activity has been of paramount concern to cognitive neuroscientists over the past few decades (Dixon et al., 2014). Although this work has been, and remains, critical in advancing our understanding of brain function, its dominance has led to the relative neglect among researchers of self-generated forms of cognition. To continue without redressing this imbalance in our scientific focus would be to ignore some of the most remarkable and complex processes of which our brains are capable. A deeper investigation of more fleeting, spontaneous forms of cognition will be necessary to move cognitive neuroscience toward a neurobiological understanding of higher mental functions like creativity and imagination. To paraphrase William James (James, 1985): No account of cognition in its totality can be final which leaves these other forms of consciousness quite disregarded.
CHAPTER 3 – STUDY 2: NEURAL CORRELATES OF SELF-GENERATED THOUGHT IN SLEEP AND DREAMING: A PRELIMINARY META-ANALYTIC ASSESSMENT OF FUNCTIONAL NEUROIMAGING INVESTIGATIONS

Introduction: the multiplicity of sleep

Awareness of the subjective multiplicity of sleep goes back thousands of years – at least as far as the ancient Indian philosophical texts known as the *Upanishads* composed around the 6th century B.C.E. (Hume, 1921, Prabhavananda et al., 1984, Deutsch and Dalvi, 2004, Sharma, 2012). Ancient Indian philosophers clearly recognized a distinction between dreamless sleep, dreaming, and even ‘lucid’ dreaming – being aware that one is dreaming while dreaming (Prabhavananda et al., 1984, Sharma, 2012). In the West, Aristotle made strikingly prescient observations for his time: he recognized both dreamless and dreaming sleep; described what we today call sleep-onset hypnagogic imagery; correctly hypothesized that dreaming represents the activity of our perceptual faculties in the absence of external inputs; and even recognized the possibility of lucid dreaming (Aristoteles and Gallop, 1996, Barbera, 2008). And at least one thousand years ago, Tibetan Buddhist practitioners had developed sophisticated cognitive practices oriented toward increasing metacognitive awareness during dreamless sleep and dreaming (Gillespie, 1988, Wangyal, 1998). All of these traditions began a fruitful mapping of quantitative and qualitative psychological (Ψ) differences in various mental states throughout the sleep cycle.

Western science began finally to put these observations on a firmer footing in the mid-20th century, with the discovery that surface-recorded brain electrical potentials could dissociate between several sleep stages (Aserinsky and Kleitman, 1953, 1955, Dement and Kleitman,
1957a, b, Monroe et al., 1965). This research, the first to definitively identify neurobiological (Φ) markers related to particular cognitive states (ψ) during sleep, led to the well-known classification of sleep into rapid-eye-movement (REM) and non-rapid-eye-movement (NREM) stages (Rechtschaffen and Kales, 1968), with four major stages generally recognized by contemporary researchers (NREM1, 2, 3/4, and REM).

These stages have not been equally recognized or researched over the past few decades. REM and NREM (which had not yet been differentiated into substages, and was collectively known as ‘slow wave sleep’ – SWS) were intensively investigated from the beginning of modern sleep and dream science in the 1950’s. It was rapidly recognized that these stages were characterized by differences in subjective experience – most notably, by the high frequency of dream reports following awakening from REM, but the relative paucity of such reports following awakenings from NREM (reviewed by Nielsen, 1999, 2000).

A pair of more marginal and difficult to investigate stages were largely ignored until recently: NREM1 (sleep onset) and so-called ‘lucid dreaming,’ in which one is aware of the fact that one is dreaming, while dreaming (LaBerge et al., 1981). Detailed investigation of the electrophysiological substages and phenomenological content of sleep onset (NREM1), although begun in the 1960’s (Foulkes and Vogel, 1965, Foulkes et al., 1966, Vogel et al., 1966), was not conducted in earnest until the 1990’s (Hori et al., 1994, Tanaka et al., 1996, 1997, Hayashi et al., 1999). Lucid REM sleep dreaming (LREM) still remains controversial to many researchers; pioneering but tenuous polysomnographic research from the 1980’s (LaBerge, 1980, LaBerge et al., 1981, Fenwick et al., 1984, LaBerge et al., 1986) has continued to be replicated and extended, however (LaBerge and Levitan, 1995, Holzinger et al., 2006, Voss et al., 2009), as
well as investigated with more sophisticated methods, such as combined EEG-fMRI (Dresler et al., 2011, Dresler et al., 2012) and transcranial direct current stimulation (Stumbrys et al., 2013).

**Sleep mentation as self-generated thought**

We have argued at length elsewhere that mentation during sleep, particularly dreaming *per se*, can be viewed as an intensified form of waking self-generated thought (Fox et al., 2013, Domhoff and Fox, 2015). The basis for this claim is twofold: both the subjective experience of dreaming, and its neurophysiological correlates (as indexed by REM sleep), strongly parallel those of waking mind-wandering and related forms of self-generated thought.

Waking self-generated thought is typically characterized by auditory and visual imagery, ubiquitous affect, a strong focus on current concerns and social interactions, and varying degrees of narrative structure (Fox et al., 2013, Andrews-Hanna et al., 2014, Fox et al., 2015; see also Chapter 1). The same statements can be made about REM sleep mentation, with the qualification that these characteristics in fact tend to be heightened or exaggerated in dreaming: the audiovisual world is fully present and immersive, emotions more intense and perhaps more ubiquitous, social characters more numerous and interactions with them more elaborate, and narrative structure is extended over time and in more complex ways (Windt, 2010, Fox et al., 2013, Domhoff and Fox, 2015).

A similar parallel is observed at the neurophysiological level. Waking self-generated thought, as compared to active focus on a task or external stimulus, is associated with a relatively consistent pattern of brain activations centered on the default network and extending into medial occipital areas involved in visual imagery, as well as some executive brain regions (see Chapter 2, above). When we previously meta-analyzed functional neuroimaging studies of REM sleep,
during which dreaming occurs approximately 80% of the time (Hobson et al., 2000), we found that many of the same brain areas implicated in waking self-generated thought were even more strongly recruited during REM sleep, including medial prefrontal cortex, numerous medial temporal lobe structures, and medial occipital areas (Fox et al., 2013). Additionally, by slightly relaxing statistical thresholds, further overlap in the inferior parietal lobule, another key default network region, was revealed (Domhoff and Fox, 2015).

Overall, these results convinced us that dreaming, and its most common neurophysiological correlate, REM sleep, show an overall intensification or amplification of both the subjective qualities and neural recruitment associated with waking self-generated cognition (cf. Fig. 3 in Fox et al., 2013). Because the NREM sleep stages are also characterized by variable levels of cognitive activity and dream experience, determining their general neural correlates is an attractive target for research, because a general understanding of these neural substrates would allow further examination of the hypothesis that self-generated thought has a common brain basis, regardless of the particular conscious state in which it takes place. Determining the general neural correlates of the NREM sleep stages was therefore the main aim of this preliminary meta-analysis.

Sleep can be meaningfully dissociated into stages

The preceding overview only hints at the enormous body of work that has been conducted over the past 60 years within a paradigm whose core assumption is that sleep stages can be meaningfully dissociated and more-or-less independently investigated. Loomis and colleagues (Loomis et al., 1935, Loomis et al., 1937) were the first to provide detailed descriptions of distinct neurophysiological stages in normal human sleep, and much subsequent
work has followed, corroborated, and expanded on these efforts. What are the criteria upon which these sleep stages are distinguished – and are they valid?

Nielsen (2014) rightly points out that the widespread use of standard sleep stage scoring criteria (Rechtschaffen and Kales, 1968) has led to an artificially categorical view of sleep stages, accompanied by tacit assumptions of both mutual exclusivity and abrupt transitions. Even the narrow use of just a few electrophysiological markers cannot support such a view; and of the hundreds of potential physiological and neural markers that fluctuate throughout the sleep-wake cycle, only a select few are routinely employed to differentiate sleep stages (Nielsen, 2014). These facts should put us on guard against any facile reification of distinct sleep stages.

A key question therefore needs to be asked: Are sleep stages a fact of neurophysiology or an investigative convenience? The answer, we believe, is that they are somewhere in between. While keeping the above caveats firmly in mind and refraining from reifying classification schemes as actual entities, we agree with the conclusions of various comprehensive reviews of this issue (e.g., (Hobson et al., 2000, Nielsen, 2000): persuasive evidence argues for the distinctiveness of sleep stages in general. Although the various major sleep stages share features in common, can oscillate back and forth unpredictably, and may hybridize and give rise to not-easily-classified transitional stages (Nielsen, 2014), meaningful statements can nonetheless be made about their characteristic patterns of phenomenology, electrophysiology, and neurochemistry (Hobson et al., 2000, Nielsen, 2000). In the following sections we present these multiple lines of evidence in support of the utility and plausibility of distinctive (if not entirely mutually exclusive) sleep stages. We argue that these neurophysiological and phenomenological idiosyncrasies lead to the strong and testable hypothesis that patterns of brain activation, as measured by relatively non-invasive functional neuroimaging methods like fMRI and positron
emission tomography (PET), should vary accordingly across the NREM and REM sleep stages. Moreover, the finding that REM sleep (with high chances of dreaming) shares many neural correlates with waking self-generated thought (Domhoff, 2011, Fox et al., 2013, Domhoff and Fox, 2015), coupled with the knowledge that self-generated thought frequency and vividness differ markedly across sleep stages (Hobson et al., 2000, Nielsen, 2000), leads to the even stronger hypothesis that specific brain areas involved in self-generated thought should show corresponding activation and deactivation throughout the sleep cycle in concert with subjectively-experienced differences in content reports.

The body of this chapter provides a systematic and quantitative meta-analytic treatment of all neuroimaging studies of sleep and dreaming in humans, in an effort to synthesize what has been learned from two decades of investigations of brain activations and deactivations throughout the sleep cycle. The aim is not to argue for strict one-to-one isomorphisms between mental states ($\psi$) and neuromarkers ($\Phi$), but rather to summarize the current evidence for broad and intriguing correspondences between first- and third-person levels of description.

**Phenomenological dissociation of sleep stages**

The subjectively-experienced (‘phenomenological’ for our purposes) differences in the experience of sleep stages have been noted for millennia (Aristoteles and Gallop, 1996) (Deutsch and Dalvi, 2004, Sharma, 2012, Thompson, 2014), but systematic research using large, representative samples has taken place mostly in the past few decades (Nielsen, 2000, Domhoff, 2003). Although many methodological difficulties (and almost as many theoretical deadlocks) have burdened this otherwise burgeoning field, some general conclusions can be cautiously
drawn regarding differences in the frequency, quality, and content of mentation across sleep stages.

Nielsen (1999; 2000) has thoroughly summarized this literature, highlighting the critical distinction between cognitive activity in general (which can include thought-like mentation, isolated flashes of imagery, and so on) and truly immersive and hallucinatory dreaming proper (a particular subset of cognitive activity). Whereas cognitive activity is fairly prevalent throughout all sleep stages (at least 40% of awakenings from any given sleep stage will lead to a report of some kind of cognitive activity; Nielsen, 1999), dreaming proper is largely restricted to REM sleep and certain sub-stages of NREM1 sleep onset, and is comparatively rare during other NREM sleep stages (Table 4).

### Table 4. Frequency of subjective reports of cognitive activity and dreaming across sleep stages.

<table>
<thead>
<tr>
<th>Sleep Stage</th>
<th>Cognitive Activity</th>
<th>Dreaming</th>
<th>Predominant EEG rhythm</th>
</tr>
</thead>
<tbody>
<tr>
<td>NREM1 (sleep onset)</td>
<td>~40%</td>
<td>~35%</td>
<td>Alpha, Theta</td>
</tr>
<tr>
<td>NREM2</td>
<td>50%</td>
<td>~15-25%</td>
<td>Theta, Spindles (Beta)</td>
</tr>
<tr>
<td>NREM3/4 (SWS)</td>
<td>40-50%</td>
<td>~10%</td>
<td>Delta</td>
</tr>
<tr>
<td>REM</td>
<td>80%</td>
<td>80%</td>
<td>Theta, Beta</td>
</tr>
<tr>
<td>LREM</td>
<td>100%</td>
<td>100%</td>
<td>Theta, Beta, Gamma</td>
</tr>
</tbody>
</table>

EEG: electroencephalography; LREM: lucid REM; NREM: non-rapid eye movement; REM: rapid eye movement; SO: sleep onset; SWS: slow-wave sleep. Table based upon comprehensive reviews by Nielsen (1999; 2000) and Hobson et al. (2000).

**Electrophysiological dissociation of sleep stages**

The first and best-known neurophysiological (Φ) division of sleep is based on scalp electrode recordings of pooled neuronal electrical potentials, presumed to represent primarily the summation of post-synaptic potentials throughout dendritic arbors and cell somata, and to a
lesser extent, synchronous discharge (action potential firing) of populations of neurons (Olejniczak, 2006, Buzsáki et al., 2012).

The central findings regarding EEG correlates of sleep are summarized in Table 4 and Figure 4. Briefly, resting (eyes-closed) wakefulness is characterized by alpha rhythms (8-12 Hz); the transition to NREM1 (sleep onset) is defined by the gradual disappearance of alpha and the appearance of theta (4-7 Hz) ripples and rolling eye movements. NREM2 begins when high-frequency spindles (in the beta frequency; 12.5-30 Hz) and large-amplitude K-complexes appear frequently in the EEG. NREM3/4 or SWS is characterized instead by very slow delta band (0.5-4 Hz) activity, synchronized in large-amplitude waves. Finally, REM sleep involves a return to highly desynchronized and low voltage activity predominantly in the theta and beta bands, similar to active wakefulness. Lucid REM sleep, in the few investigations so far conducted, involves an EEG pattern similar to REM sleep but with increased power in the gamma (>30 Hz) band (Voss et al., 2009, Voss et al., 2014). As these electrophysiological differences are almost universally accepted, well-validated, and expertly reviewed elsewhere (Williams et al., 1974, Antrobus, 1991, Silber et al., 2007), they are not discussed further here.
<table>
<thead>
<tr>
<th>Stage</th>
<th>sub stages*</th>
<th>EEG Signature</th>
<th>EEG Signature Waveform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wake</td>
<td>1</td>
<td>Alpha wave train</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>Alpha wave intermittent (&gt;50%)</td>
<td></td>
</tr>
<tr>
<td>NREM 1</td>
<td>3</td>
<td>Alpha wave intermittent (&lt;50%)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>EEG flattening (&lt;20 μV)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Theta ripples</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Vertex sharp wave (&lt;200 μV)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>&gt; 1 Vertex sharp wave (&lt;200 μV)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Incomplete spindle</td>
<td></td>
</tr>
<tr>
<td>NREM 2</td>
<td>9</td>
<td>Complete Spindle</td>
<td></td>
</tr>
<tr>
<td>NREM 3-4</td>
<td></td>
<td>Delta</td>
<td></td>
</tr>
<tr>
<td>REM</td>
<td></td>
<td>Theta</td>
<td></td>
</tr>
</tbody>
</table>

*Figure 4. Main electrophysiological correlates of each sleep stage. All NREM1 sub-stages are included, as is eyes-closed waking rest, for comparison. Note the easily-differentiable EEG signature accompanying each sleep stage. Reproduced from Stenstrom, Fox, et al. (2012).
Neurochemical dissociation of sleep stages

The patterns of neurochemical activity throughout the sleep-wake cycle are exceedingly complex and consequently poorly understood. At least a dozen major neurotransmitters are involved in regulating NREM and REM, by virtue of either increased or decreased activity (compared to waking) during various sleep stages, but many secondary players with less clear roles are also involved (Kahn et al., 1997, Gottesmann, 1999, Hobson et al., 2000, Pace-Schott and Hobson, 2002, Lena et al., 2005, Stenberg, 2007). Moreover, changes in neurotransmitter levels are far from a global phenomenon throughout the brain: region-by-transmitter interactions have in some cases been experimentally demonstrated (e.g., for dopamine: (Lena et al., 2005)), and it seems likely that activity levels for other neurotransmitters will also vary based not just on sleep stage, but also which region of the brain is being investigated. These complexities are further exacerbated by the existence of many receptor subtypes for each neurotransmitter, and the concomitant (and often unknown – at least for sleep) differences in downstream effects caused by the actions of a single neurotransmitter (Monti and Monti, 2007). A final and major difficulty is that implanting recording electrodes directly into the subcortical nuclei responsible for manufacturing and/or disseminating these neurotransmitters is the ideal method of investigation – an approach typically precluded in humans. Almost everything that is known about the neurochemistry of sleep has therefore been drawn from studies of animals with variable phylogenetic proximity to humans, such as rats, cats, rabbits, and monkeys (Jones, 1991, Gottesmann, 1999, Jones, 2005, Lena et al., 2005, Stenberg, 2007).

The enormous difficulty of studying the neurochemistry of sleep and dreaming have so far precluded the formulation of a clear model of each neurotransmitter’s relative activity throughout NREM and REM sleep stages, much less what the functional implications of such neurochemical heterogeneity might be. Nonetheless, decades of research have yielded some
broad trends that appear to be generally reliable, and are bolstered by converging evidence across species, brain regions, and methods. We summarize these general trends in Table 5.

Although a stage-by-stage model is premature given the current limits of our knowledge, broad trends easily distinguish waking at least from REM and NREM sleep. Generally speaking, all major neurotransmitters show some level of tonic activity during waking; conversely, all of these neurotransmitters show a greater or lesser decrease of activity during various NREM stages of sleep (it should be noted, however, that much of this data is derived only from the later stages of ‘slow-wave’ NREM sleep). Finally, REM sleep shows an intermediate pattern: most neurotransmitter activity is decreased relative to waking, but notably, acetylcholine and dopamine levels are elevated (Table 5).
Table 5. Neurochemical profiles of the various stages of sleep as compared to waking.

<table>
<thead>
<tr>
<th>Neurotransmitter</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>REM (↑ Dreaming)</td>
</tr>
<tr>
<td>ACh</td>
<td>↑↑</td>
</tr>
<tr>
<td>Asp</td>
<td>↓</td>
</tr>
<tr>
<td>DA</td>
<td>↑↑</td>
</tr>
<tr>
<td>GABA</td>
<td>?</td>
</tr>
<tr>
<td>Glu</td>
<td>↓↓</td>
</tr>
<tr>
<td>HA</td>
<td>↓↓</td>
</tr>
<tr>
<td>NE</td>
<td>↓↓</td>
</tr>
<tr>
<td>5-HT</td>
<td>↓↓</td>
</tr>
</tbody>
</table>

ACh: acetylcholine; Asp: aspartate; DA: dopamine; GABA: \(\gamma\)-amino butyric acid; Glu: glutamate; HA: histamine; NE: norepinephrine/noradrenaline; NREM: non-rapid-eye-movement sleep; REM: rapid-eye-movement sleep; 5-HT: serotonin; SWS: slow wave sleep. Based on data from (Gottesmann, 1999, Pace-Schott and Hobson, 2002, Lena et al., 2005, Hobson, 2009).

Functional neuroimaging of sleep and dreaming: a fourth line of evidence

The preceding sections have highlighted some of the key evidence supporting the hypothesis that stages of sleep are readily dissociable in terms of their (i) phenomenology, (ii) electrophysiology, and more provisionally, (iii) neurochemistry. These many heterogeneities across sleep stages lead to the strong and testable hypothesis that brain activations and deactivations, as detected with relatively non-invasive functional neuroimaging modalities, might also show dissociable patterns throughout the sleep cycle. Functional neuroimaging therefore provides a fourth line of evidence that could either corroborate (or possibly contradict) the ample body of research suggesting phenomenological and neurophysiological dissociability. As noted above, the phenomenological variations in self-generated thought across sleep stages further suggest that differential activation patterns, if they exist, should exhibit some relationship
with the numerous regions implicated in self-generated thought in waking (Fox et al., 2015) and REM sleep (Fox et al., 2013).

**Aims of the present meta-analysis**

Despite a profusion of neuroimaging studies elucidating the neural basis of sleep and dreaming (Pace-Schott, 2007, Maquet, 2010, Desseilles et al., 2011), no meta-analysis providing a comprehensive overview has yet been conducted. The increasing number and wide diversity of reported brain activations and deactivations observed across different sleep stages, however, demands a meta-analytic synthesis. The drastically different subjective content reported from laboratory awakenings across the various sleep stages corroborates the distinctive patterns of observed brain activity, but there has yet to be any systematic attempt to relate these differing first-person experiences to particular patterns of brain activation. Recently, we conducted such a meta-analysis and review of subjective content for REM sleep and dreaming (Fox et al., 2013); here, we expand upon those results to include functional neuroimaging and subjective report data from all sleep stages. Beyond being the first quantitative meta-analysis of functional neuroimaging studies of sleep and dreaming, this is the first review to attempt a systematic relating of subjective experience, electrophysiology, neurochemistry, and brain blood-flow-related activity across each sleep stage.
Table 6. Brain states investigated, with number of contributing studies and foci, and number of meta-analytic results obtained.

<table>
<thead>
<tr>
<th>Cognitive / Brain State</th>
<th>Contributing Studies (N)</th>
<th>Contributing Foci (A/D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypnagogia / NREM1</td>
<td>4</td>
<td>27/47</td>
</tr>
<tr>
<td>Light Sleep / NREM2</td>
<td>4</td>
<td>17/32</td>
</tr>
<tr>
<td>Deep Sleep / NREM3/4 (SWS)</td>
<td>7</td>
<td>30/150</td>
</tr>
<tr>
<td>Dreaming / REM</td>
<td>8</td>
<td>84/49</td>
</tr>
<tr>
<td>Lucid dreaming / LREM</td>
<td>1</td>
<td>10/0</td>
</tr>
</tbody>
</table>

A: activations; D: deactivations; LREM: lucid REM; NREM: non-REM; REM: rapid eye movement; SWS: slow wave sleep.

Methods

Study selection

Three experimenters scoured Google Scholar, PubMed, and PsycInfo online databases for any study whose title or abstract included keywords such as ‘PET,’ ‘fMRI,’ and ‘neuroimaging,’ in combination with keywords such as ‘sleep,’ ‘dreaming,’ ‘NREM,’ or ‘REM.’ The reference list of each study found was also consulted, as were the bibliographies of several major review of functional neuroimaging of sleep (Hobson et al., 2000)

Our search yielded 58 functional neuroimaging (PET or fMRI) studies of REM and NREM sleep. As inclusion criteria, we required that studies employ a baseline of resting wakefulness (either pre- or post-sleep), in order to minimize the confounding effects of various tasks and baseline conditions; that studies involve neurological and psychiatrically healthy participants; and that studies reported peak coordinates for contrasts of interest in standard stereotactic space (either Talairach or MNI coordinates). A total of 16 studies met these inclusion
Because many studies examined more than one sleep stage, a total of 24 contrasts of interest were obtained from this pool of studies.

Besides failing to meet our basic inclusion criteria, reasons for exclusion included the addition of extraneous factors (e.g., auditory stimulation during sleep), inclusion of clinical populations or pharmacological agents, failure to provide information for peak foci of activation, or lack of an appropriate baseline (e.g., studies comparing REM sleep during phasic rapid-eye-movement events with regular tonic REM sleep). Note that many studies employed a sleep deprivation paradigm to aid participants in sleeping in the scanner (see discussion of limitations, below).

**Table 7.** Studies included in meta-analyses of sleep and dreaming.

<table>
<thead>
<tr>
<th>Study</th>
<th>Modality</th>
<th>N</th>
<th>Stage(s) investigated</th>
<th>Sleep deprivation?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maquet et al. (1996)</td>
<td>PET</td>
<td>11</td>
<td>REM</td>
<td>Y</td>
</tr>
<tr>
<td>Braun et al. (1997)</td>
<td>PET</td>
<td>37</td>
<td>SWS, REM</td>
<td>Y</td>
</tr>
<tr>
<td>Maquet et al. (1997)</td>
<td>PET</td>
<td>11</td>
<td>SWS</td>
<td>Y</td>
</tr>
<tr>
<td>Nofzinger et al. (1997)</td>
<td>PET</td>
<td>6</td>
<td>REM</td>
<td>N</td>
</tr>
<tr>
<td>Braun et al. (1998)</td>
<td>PET</td>
<td>10</td>
<td>REM</td>
<td>Y</td>
</tr>
<tr>
<td>Kajimura et al. (1999)</td>
<td>PET</td>
<td>18</td>
<td>SWS</td>
<td>Y</td>
</tr>
<tr>
<td>Finelli et al. (2000)</td>
<td>PET</td>
<td>8</td>
<td>REM</td>
<td>Y</td>
</tr>
<tr>
<td>Maquet et al. (2000)</td>
<td>PET</td>
<td>5</td>
<td>REM</td>
<td>N</td>
</tr>
<tr>
<td>Peigneux et al. (2001)</td>
<td>PET</td>
<td>12</td>
<td>REM</td>
<td>N</td>
</tr>
<tr>
<td>Balkin et al. (2002)</td>
<td>PET</td>
<td>27</td>
<td>NREM2</td>
<td>Y</td>
</tr>
<tr>
<td>Kjaer et al. (2002)</td>
<td>PET</td>
<td>8</td>
<td>NREM1</td>
<td>N</td>
</tr>
<tr>
<td>Maquet et al. (2005)</td>
<td>PET</td>
<td>22</td>
<td>SWS, REM</td>
<td>N</td>
</tr>
<tr>
<td>Kaufmann et al. (2006)</td>
<td>fMRI</td>
<td>9</td>
<td>NREM1, NREM2, SWS</td>
<td>Y</td>
</tr>
<tr>
<td>Picchioni et al. (2008)</td>
<td>fMRI</td>
<td>4</td>
<td>NREM1</td>
<td>N</td>
</tr>
<tr>
<td>Andrade et al. (2011)</td>
<td>fMRI</td>
<td>25</td>
<td>NREM1, NREM2, SWS</td>
<td>N</td>
</tr>
<tr>
<td>Koike et al. (2011)</td>
<td>fMRI</td>
<td>12</td>
<td>NREM2, SWS, REM</td>
<td>N</td>
</tr>
</tbody>
</table>
**Activation likelihood estimation (ALE) meta-analysis**

We used a quantitative, random-effects meta-analytic method known as activation likelihood estimation (ALE) (Turkeltaub et al., 2002, Laird et al., 2005, Eickhoff et al., 2009, Eickhoff et al., 2012) implemented in the software program GingerALE 2.3.1 (San Antonio, TX: UT Health Science Center Research Imaging Institute). The most recent ALE algorithm tests for above-chance clustering of peak foci from different experiments included in the meta-analysis (Eickhoff et al., 2009, Eickhoff et al., 2012) by comparing actual activation foci locations/clustering with a null distribution created by distributing the same number of foci randomly throughout the brain, through several thousand iterations. Included activation foci were smoothed using a full-width half maximum (FWHM) Gaussian kernel dependent on the sample size (subjects) of the experiment from which foci were drawn (larger sample -> smaller smoothing kernel – empirically determined by (Eickhoff et al., 2009, Eickhoff et al., 2012)). Resulting statistical maps show clusters where convergence between activation foci is greater than would be expected by chance (i.e., if foci from each experiment were distributed independently).

Statistical maps were thresholded using a false discovery rate (FDR – (Genovese et al., 2002)) of \( q = .10 \) (pN) and a cluster threshold of \( k = 100 \text{ mm}^3 \). We used these relatively liberal height and extent thresholds due to the very small number of studies available for meta-analysis. To display results, we used template brain images from GingerALE 2.3.1 displayed in the ‘Mango’ software package (San Antonio, TX: UT Health Science Center Research Imaging Institute). Final figures were created using MRICron.
**Reporting of meta-analytic results**

Though we recently conducted a similar meta-analysis of neuroimaging studies of REM sleep (Fox et al., 2013), we update and expand that analysis here. We conducted a total of 8 separate meta-analyses: meta-analytic assessments of both whole-brain activations and deactivations for each of four sleep stages (NREM1, NREM2, NREM3/4 [SWS], and REM). In addition, the results of a single fMRI investigation of lucid dreaming (Dresler et al., 2012) were consulted for comparison purposes. Activation and deactivation maps were the concatenated onto a single brain template for visualization purposes.

**Results**

**Neural correlates of NREM1 sleep**

For NREM1 sleep, the most salient finding was widespread visual network activations, including clusters in the fusiform gyrus bilaterally (BA 20/36/37), the middle occipital gyrus (BA 19), and the cuneus (BA 18). Deactivations were most prominent in the thalamus bilaterally, in somatomotor cortex (BA 4), in premotor cortex bilaterally (BA 6), and in right dorsolateral prefrontal cortex (BA 9).
Table 8. Activations and deactivations associated with NREM1 sleep.

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster Size (mm$^3$)</th>
<th>Side</th>
<th>Peak Coordinates (x, y, z)</th>
<th>Peak ALE value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusiform gyrus (BA 20/36)</td>
<td>152</td>
<td>R</td>
<td>44, -38, -24</td>
<td>0.0089</td>
</tr>
<tr>
<td>Fusiform gyrus (BA 37)</td>
<td>152</td>
<td>L</td>
<td>-22, -44, -18</td>
<td>0.0089</td>
</tr>
<tr>
<td>Internal capsule</td>
<td>152</td>
<td>L</td>
<td>-26, -24, 0</td>
<td>0.0076</td>
</tr>
<tr>
<td>Superior temporal gyrus (BA 22)</td>
<td>152</td>
<td>L</td>
<td>-52, -28, 2</td>
<td>0.0089</td>
</tr>
<tr>
<td>Middle occipital gyrus (BA 19)</td>
<td>152</td>
<td>R</td>
<td>42, -68, 8</td>
<td>0.0089</td>
</tr>
<tr>
<td>Cuneus (BA 18)</td>
<td>152</td>
<td>R</td>
<td>16, -82, 30</td>
<td>0.0089</td>
</tr>
<tr>
<td>Premotor cortex (BA 6)</td>
<td>152</td>
<td>R</td>
<td>28, 10, 48</td>
<td>0.0076</td>
</tr>
<tr>
<td>Precuneus (BA 7)</td>
<td>112</td>
<td>L</td>
<td>-18, -66, 40</td>
<td>0.0073</td>
</tr>
<tr>
<td><strong>Deactivations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thalamus</td>
<td>1240</td>
<td>R</td>
<td>4, -20, 6</td>
<td>0.0116</td>
</tr>
<tr>
<td>Somatomotor cortex (BA 4)</td>
<td>448</td>
<td>L</td>
<td>-45, -8, 52</td>
<td>0.0081</td>
</tr>
<tr>
<td>Premotor cortex (BA 6)</td>
<td>320</td>
<td>L</td>
<td>-36, 6, 40</td>
<td>0.0073</td>
</tr>
<tr>
<td>Thalamus</td>
<td>280</td>
<td>L</td>
<td>-6, -14, 16</td>
<td>0.0067</td>
</tr>
<tr>
<td>Premotor cortex (BA 6)</td>
<td>184</td>
<td>R</td>
<td>14, 8, 56</td>
<td>0.0078</td>
</tr>
<tr>
<td>Dorsolateral prefrontal cortex (BA 9)</td>
<td>160</td>
<td>R</td>
<td>31, 23, 28</td>
<td>0.0071</td>
</tr>
<tr>
<td>Inferior parietal lobule (BA 39/7)</td>
<td>160</td>
<td>L</td>
<td>-31, -56, 44</td>
<td>0.0073</td>
</tr>
<tr>
<td>Precuneus (BA 7)</td>
<td>160</td>
<td>L</td>
<td>-14, -63, 46</td>
<td>0.0073</td>
</tr>
</tbody>
</table>

**Neural correlates of NREM2 sleep**

NREM2 was associated with marginal results overall, consistent with the very small sample of studies included in this meta-analysis. Activations included some default network (BA 32) and visual network (BA 18) regions, as well as clusters in the cerebellum and caudate nucleus. Deactivations were present in the inferior parietal lobule bilaterally (BA 40), dorsal anterior cingulate cortex (BA 24), and the cerebellum, among other regions (see Table 9).
**Table 9.** Activations and deactivations associated with NREM2 sleep.

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster Size (mm$^3$)</th>
<th>Side</th>
<th>Peak Coordinates (x, y, z)</th>
<th>Peak ALE value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebellum</td>
<td>160</td>
<td>M</td>
<td>-2, -77, -28</td>
<td>0.0087</td>
</tr>
<tr>
<td>Caudate nucleus (body)</td>
<td>160</td>
<td>R</td>
<td>12, 6, 9</td>
<td>0.0087</td>
</tr>
<tr>
<td>Anterior cingulate cortex (BA 32)</td>
<td>160</td>
<td>M</td>
<td>-8, 28, 19</td>
<td>0.0087</td>
</tr>
<tr>
<td>Lingual gyrus (BA 18)</td>
<td>152</td>
<td>M</td>
<td>8, -88, -8</td>
<td>0.0089</td>
</tr>
<tr>
<td>Superior temporal gyrus (BA 21/22)</td>
<td>152</td>
<td>L</td>
<td>-64, -8, -4</td>
<td>0.0089</td>
</tr>
<tr>
<td>Posterior mid-cingulate cortex (BA 24)</td>
<td>152</td>
<td>M</td>
<td>6, -12, 36</td>
<td>0.0089</td>
</tr>
<tr>
<td>Somatomotor cortex (BA 4)</td>
<td>136</td>
<td>L</td>
<td>-46, -6, 56</td>
<td>0.0089</td>
</tr>
<tr>
<td>Superior temporal gyrus (BA 42)</td>
<td>128</td>
<td>R</td>
<td>68, -18, 8</td>
<td>0.0089</td>
</tr>
<tr>
<td><strong>Deactivations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebellum</td>
<td>152</td>
<td>M</td>
<td>-6, -68, -34</td>
<td>0.0071</td>
</tr>
<tr>
<td>Inferior temporal gyrus (BA 20/37)</td>
<td>152</td>
<td>L</td>
<td>-56, -62, -12</td>
<td>0.0071</td>
</tr>
<tr>
<td>Inferior parietal lobule (BA 40)</td>
<td>152</td>
<td>R</td>
<td>46, -30, 48</td>
<td>0.0071</td>
</tr>
<tr>
<td>Precuneus (BA 7)</td>
<td>152</td>
<td>R</td>
<td>14, -50, 60</td>
<td>0.0071</td>
</tr>
<tr>
<td>Premotor cortex (BA 6)</td>
<td>152</td>
<td>R</td>
<td>28, 6, 62</td>
<td>0.0071</td>
</tr>
<tr>
<td>Inferior parietal lobule (BA 40)</td>
<td>144</td>
<td>L</td>
<td>-48, -28, 60</td>
<td>0.0071</td>
</tr>
<tr>
<td>Dorsal anterior cingulate cortex (BA 24)</td>
<td>128</td>
<td>M</td>
<td>-8, 5, 46</td>
<td>0.0058</td>
</tr>
</tbody>
</table>
Neural correlates of NREM3/4 sleep (SWS)

SWS was associated with modest clusters of activation in the superior temporal gyrus bilaterally (BA 22), in right parahippocampal cortex (BA 19), and in various visual network regions (BA 18/19). In contrast to these small clusters of activation, massive deactivations were observed throughout the brain, most notably in the thalamus, and in prefrontal executive regions, including orbitofrontal cortex bilaterally (BA 47), right rostrolateral prefrontal cortex (BA 10), left ventrolateral prefrontal cortex (BA 47), and dorsal anterior cingulate cortex (BA 32). Further, deactivations were observed in default network regions, including posterior cingulate cortex (BA 31) and multiple clusters in medial prefrontal cortex (BA 32).
**Table 10.** Activations and deactivations associated with NREM3/4 sleep (SWS).

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster Size (mm$^3$)</th>
<th>Side</th>
<th>Peak Coordinates (x, y, z)</th>
<th>Peak ALE value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superior temporal gyrus (BA 22)</td>
<td>256</td>
<td>L</td>
<td>-56, -23, 2</td>
<td>0.0077</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>160</td>
<td>M</td>
<td>2, -58, -18</td>
<td>0.0081</td>
</tr>
<tr>
<td>Parahippocampal cortex (BA 19)</td>
<td>160</td>
<td>R</td>
<td>20, -59, -13</td>
<td>0.0079</td>
</tr>
<tr>
<td>Middle temporal gyrus (BA 37)</td>
<td>160</td>
<td>L</td>
<td>-46, -49, -8</td>
<td>0.0081</td>
</tr>
<tr>
<td>Superior temporal gyrus (BA 22)</td>
<td>160</td>
<td>L</td>
<td>-46, -34, -2</td>
<td>0.0082</td>
</tr>
<tr>
<td>Lingual gyrus (BA 18)</td>
<td>160</td>
<td>M</td>
<td>10, -85, 0</td>
<td>0.0081</td>
</tr>
<tr>
<td>Posterior insula (BA 13)</td>
<td>160</td>
<td>L</td>
<td>-40, -23, 2</td>
<td>0.0079</td>
</tr>
<tr>
<td>Middle occipital gyrus (BA 19)</td>
<td>160</td>
<td>L</td>
<td>-36, -78, 10</td>
<td>0.0079</td>
</tr>
<tr>
<td>Lingual gyrus (BA 18)</td>
<td>152</td>
<td>M</td>
<td>10, -60, 6</td>
<td>0.0084</td>
</tr>
<tr>
<td><strong>Deactivations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thalamus</td>
<td>2336</td>
<td>L</td>
<td>-6, -22, 2</td>
<td>0.0180</td>
</tr>
<tr>
<td>Thalamus/hypothalamus/subgenual anterior cingulate cortex (BA 25)</td>
<td>1312</td>
<td>M</td>
<td>0, -6, -6</td>
<td>0.0150</td>
</tr>
<tr>
<td>Lateral globus pallidus/putamen</td>
<td>776</td>
<td>R</td>
<td>16, 8, -6</td>
<td>0.0137</td>
</tr>
<tr>
<td>Orbitofrontal cortex (BA 47)</td>
<td>536</td>
<td>L</td>
<td>-20, 20, -22</td>
<td>0.0147</td>
</tr>
<tr>
<td>Caudate nucleus (body)</td>
<td>528</td>
<td>R</td>
<td>14, 6, 8</td>
<td>0.0141</td>
</tr>
<tr>
<td>Orbitofrontal cortex (BA 47)</td>
<td>504</td>
<td>R</td>
<td>20, 24, -18</td>
<td>0.0133</td>
</tr>
<tr>
<td>Rostrolateral prefrontal cortex (BA 10)</td>
<td>504</td>
<td>R</td>
<td>26, 52, -12</td>
<td>0.0127</td>
</tr>
<tr>
<td>Medial prefrontal cortex (BA 32)</td>
<td>504</td>
<td>M</td>
<td>2, 40, 10</td>
<td>0.0134</td>
</tr>
<tr>
<td>Anterior insula (BA 13)</td>
<td>488</td>
<td>R</td>
<td>44, 16, -10</td>
<td>0.0116</td>
</tr>
<tr>
<td>Dorsal anterior cingulate cortex (BA 32)</td>
<td>464</td>
<td>M</td>
<td>-4, 22, 30</td>
<td>0.0125</td>
</tr>
<tr>
<td>Dorsal pons</td>
<td>456</td>
<td>M</td>
<td>-6, -34, -36</td>
<td>0.0117</td>
</tr>
<tr>
<td>Mid-insula (BA 13)</td>
<td>456</td>
<td>R</td>
<td>38, 6, 2</td>
<td>0.0121</td>
</tr>
<tr>
<td>Posterior cingulate cortex (BA 31)</td>
<td>432</td>
<td>M</td>
<td>-2, -34, 34</td>
<td>0.0120</td>
</tr>
<tr>
<td>Medial prefrontal cortex (BA 32)</td>
<td>392</td>
<td>M</td>
<td>10, 48, -12</td>
<td>0.0116</td>
</tr>
<tr>
<td>Lateral globus pallidus/putamen</td>
<td>296</td>
<td>L</td>
<td>-24, -4, -6</td>
<td>0.0099</td>
</tr>
<tr>
<td>Caudate nucleus (body)</td>
<td>176</td>
<td>L</td>
<td>-16, 4, 10</td>
<td>0.0087</td>
</tr>
<tr>
<td>Ventrolateral prefrontal cortex (BA 47)</td>
<td>104</td>
<td>L</td>
<td>-32, 32, -18</td>
<td>0.0091</td>
</tr>
<tr>
<td>Rostrolateral prefrontal cortex (BA 10)</td>
<td>104</td>
<td>R</td>
<td>42, 52, -6</td>
<td>0.0091</td>
</tr>
</tbody>
</table>
**Neural correlates of REM sleep**

REM sleep was associated with recruitment of widespread brain regions. Most prominent were large clusters of activation throughout the medial temporal lobe bilaterally; multiple regions within the default mode network, including two clusters in medial prefrontal cortex (BA 24 and 9/32), dorsomedial prefrontal cortex (BA 9), and orbitofrontal cortex (BA 25); and numerous visual network areas, centered on the lingual gyrus (BA 18/19). Deactivations were most salient in prefrontal executive regions, including bilateral dorsolateral prefrontal cortex (BA 9), bilateral ventrolateral prefrontal cortex (BA 47), and bilateral frontopolar cortex (BA 10). See Table 11 for full results.
Table 11. Activations and deactivations associated with REM sleep.

<table>
<thead>
<tr>
<th>Region</th>
<th>Cluster Size (mm$^3$)</th>
<th>Side</th>
<th>Peak Coordinates (x, y, z)</th>
<th>Peak ALE value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Activations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parahippocampal cortex (BA 36)</td>
<td>1120</td>
<td>R</td>
<td>24, -42, -12</td>
<td>0.0118</td>
</tr>
<tr>
<td>Caudate nucleus (body)</td>
<td>856</td>
<td>M</td>
<td>-6, 16, 12</td>
<td>0.0137</td>
</tr>
<tr>
<td>Pons/midbrain</td>
<td>848</td>
<td>M</td>
<td>8, -12, -18</td>
<td>0.0125</td>
</tr>
<tr>
<td>Parahippocampal cortex/lingual gyrus (BA 36/19)</td>
<td>832</td>
<td>L</td>
<td>-18, -50, -12</td>
<td>0.0133</td>
</tr>
<tr>
<td>Lingual gyrus (BA 19)</td>
<td>800</td>
<td>R</td>
<td>28, -68, 0</td>
<td>0.0149</td>
</tr>
<tr>
<td>Medial prefrontal cortex (BA 24)</td>
<td>704</td>
<td>M</td>
<td>4, 32, 4</td>
<td>0.0138</td>
</tr>
<tr>
<td>Paracentral lobule (BA 41)</td>
<td>640</td>
<td>M</td>
<td>0, -14, 46</td>
<td>0.0107</td>
</tr>
<tr>
<td>Parahippocampal cortex/lingual gyrus (BA 19)</td>
<td>608</td>
<td>R</td>
<td>22, -60, -10</td>
<td>0.0125</td>
</tr>
<tr>
<td>Parahippocampal cortex (BA 28/35)</td>
<td>600</td>
<td>R</td>
<td>22, -18, -16</td>
<td>0.0126</td>
</tr>
<tr>
<td>Pons</td>
<td>464</td>
<td>M</td>
<td>-2, -18, -30</td>
<td>0.0108</td>
</tr>
<tr>
<td>Parahippocampal cortex (BA 35)</td>
<td>344</td>
<td>L</td>
<td>-16, -26, -24</td>
<td>0.0094</td>
</tr>
<tr>
<td>Dorsomedial prefrontal cortex (BA 9)</td>
<td>216</td>
<td>M</td>
<td>-2, 44, 28</td>
<td>0.0096</td>
</tr>
<tr>
<td>Lingual gyrus (BA 18)</td>
<td>208</td>
<td>L</td>
<td>-10, -94, -12</td>
<td>0.0083</td>
</tr>
<tr>
<td>Orbitofrontal cortex (BA 25)</td>
<td>184</td>
<td>M</td>
<td>-2, 20, -22</td>
<td>0.0093</td>
</tr>
<tr>
<td>Medial prefrontal cortex (BA 9/32)</td>
<td>144</td>
<td>R</td>
<td>12, 44, 10</td>
<td>0.0078</td>
</tr>
<tr>
<td><strong>Deactivations</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dorsolateral prefrontal cortex (BA 9)</td>
<td>648</td>
<td>R</td>
<td>32, 44, 24</td>
<td>0.0161</td>
</tr>
<tr>
<td>Ventrolateral prefrontal cortex (BA 47)</td>
<td>592</td>
<td>R</td>
<td>38, 38, -12</td>
<td>0.0141</td>
</tr>
<tr>
<td>Posterior cingulate cortex (BA 31)/Splenium</td>
<td>520</td>
<td>L</td>
<td>-10, -34, 26</td>
<td>0.0111</td>
</tr>
<tr>
<td>Ventrolateral prefrontal cortex (BA 47)</td>
<td>504</td>
<td>L</td>
<td>-30, 40, -12</td>
<td>0.0114</td>
</tr>
<tr>
<td>Frontopolar cortex (BA 10)</td>
<td>440</td>
<td>R</td>
<td>18, 50, -16</td>
<td>0.0111</td>
</tr>
<tr>
<td>Inferior parietal lobule (BA 39)</td>
<td>368</td>
<td>R</td>
<td>56, -56, 24</td>
<td>0.0101</td>
</tr>
<tr>
<td>Inferior parietal lobule (BA 39)</td>
<td>192</td>
<td>L</td>
<td>-38, -50, 28</td>
<td>0.0095</td>
</tr>
<tr>
<td>Posterior corona radiata</td>
<td>160</td>
<td>R</td>
<td>28, -44, 20</td>
<td>0.0092</td>
</tr>
<tr>
<td>Posterior corona radiata</td>
<td>160</td>
<td>R</td>
<td>30, -52, 24</td>
<td>0.0089</td>
</tr>
<tr>
<td>Dorsolateral prefrontal cortex (BA 9)</td>
<td>160</td>
<td>R</td>
<td>44, 32, 24</td>
<td>0.0092</td>
</tr>
<tr>
<td>Region</td>
<td>Cluster Size (mm$^3$)</td>
<td>Side</td>
<td>Peak Coordinates (x, y, z)</td>
<td>Peak ALE value</td>
</tr>
<tr>
<td>--------------------------------------------</td>
<td>-----------------------</td>
<td>------</td>
<td>-----------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Inferior parietal lobule (BA 40)</td>
<td>160</td>
<td>R</td>
<td>46, -46, 24</td>
<td>0.0092</td>
</tr>
<tr>
<td>Orbitofrontal cortex (BA 11)</td>
<td>152</td>
<td>R</td>
<td>22, 32, -26</td>
<td>0.0086</td>
</tr>
<tr>
<td>Rostrolateral prefrontal cortex (BA 10)</td>
<td>152</td>
<td>L</td>
<td>-32, 54, -14</td>
<td>0.0086</td>
</tr>
<tr>
<td>Rostrolateral prefrontal cortex (BA 10)</td>
<td>152</td>
<td>L</td>
<td>-36, 56, 0</td>
<td>0.0086</td>
</tr>
<tr>
<td>Dorsolateral prefrontal cortex (BA 46)</td>
<td>152</td>
<td>L</td>
<td>-50, 40, 2</td>
<td>0.0086</td>
</tr>
<tr>
<td>Insula/inferior frontal gyrus (BA 13/44)</td>
<td>152</td>
<td>R</td>
<td>46, 18, 10</td>
<td>0.0095</td>
</tr>
<tr>
<td>Posterior corona radiata</td>
<td>152</td>
<td>R</td>
<td>28, -32, 20</td>
<td>0.0096</td>
</tr>
</tbody>
</table>
Brain activations and deactivations throughout the normal sleep cycle (panels a-l), and in lucid REM sleep dreaming (panels m-o). Frontoparietal control network regions are deactivated throughout all normal sleep stages, except lucid dreaming, consistent with the lack of cognitive control and meta-awareness in these sleep stages. Default and visual network areas, in contrast, show a stepwise relationship with sleep stage: they are most activated when rates of dreaming and other cognitive activity are highest (REM sleep and lucid REM), and deactivated when phenomenological content approaches zero (as in NREM3/4). Lucid REM sleep figure reproduced from Dresler et al. (2012).

alns: anterior insula; aIPL: anterior inferior parietal lobule; dACC: dorsal anterior cingulate cortex; DLPFC: dorsolateral prefrontal cortex; Hth: hypothalamus; IOcc: lateral occipital cortex; LTC: lateral temporal cortex; mOcc: medial occipital cortex; mPFC: medial prefrontal cortex; PCC: posterior cingulate cortex; RLPFC: rostrolateral prefrontal cortex; Thal: thalamus.

**Figure 5.** Brain activations and deactivations throughout the sleep cycle.
Discussion

As we observed stepwise relationships between subjective thought content and brain activation throughout the sleep cycle, we discuss overall trends in our results as a whole rather than focusing the discussion on one sleep stage at a time, as in the Results section. Note, however, that the putative links between subjective experiences and brain activation discussed here are almost entirely inferential and based on statistical averages of content reports generated in studies that did not use PET scans to assess brain activation. That is, as outlined in the introduction, the subjective experiences assumed to hold, on average, for each sleep stage are based on datasets entirely independent from those upon which brain activation data are based. The apparent linkages we speculate on here are therefore, at best, crude approximations of average experiential intensity and average brain activation across many different subjects; although intriguing, these correspondences need to be further explored and corroborated but much detailed research before they can be considered reliable, much less definitive. This major concern is discussed in more detail in the subsequent section on Limitations.

Default and visual network activation and their relationship to self-generated thought and imagery throughout the sleep cycle

Dreaming can be thought of as an unconstrained, hyper-associative, and highly immersive form of self-generated thought that is largely detached from external sensory inputs (Windt, 2010, Fox et al., 2013, Christoff et al., in press). We therefore expected that areas involved in memory recall and recombination, self-referential thinking, and audiovisual imagery would show heightened recruitment, compared to a restful waking baseline, in all sleep stages that have some appreciable amount of thought or dream content (especially REM sleep, but also potentially stages NREM1 and NREM2).
Generally speaking, the NREM sleep stages are associated with considerably lower rates of dreaming than REM sleep (see Introduction) – and the mentation that does occur tends to be less visuospatial and immersive, and more of a conceptually-focused inner monologue (Hobson et al., 2000, Nielsen, 2000, Nir and Tononi, 2010). During NREM1, however, highly vivid visual imagery and occasionally full-blown (if short-lived) dreaming can occur (Nielsen, 1992, Hori et al., 1994, Hayashi et al., 1999, Stenstrom et al., 2012). Consistent with these phenomenological reports, our meta-analysis showed that NREM1 was primarily associated with widespread activation of secondary and tertiary visual regions, including the fusiform gyri bilaterally, the middle occipital gyrus (BA 19), and the cuneus (BA 18). We also observed activation in primary auditory cortex (superior temporal gyrus), consistent with the high rates of auditory imagery experienced at sleep onset (Stenstrom et al., 2012). Absent, however, were any notable activations in default network regions; this absence of activation, however, can be reconciled with the typical brevity and diminished sense self (compared to other sleep stage mentation reports) that characterize sleep onset imagery (Cicogna et al., 1998).

Late in the night, when REM sleep predominates in the sleep cycle, NREM2 can give rise to a high frequency of immersive dream experiences indistinguishable from REM sleep reports (Cicogna et al., 1998). Cognitive activity of some kind is relatively frequent during early-night NREM2 (~50% of awakenings), but dreaming proper is more rare, reported from roughly 15-25% of NREM2 awakenings (Goodenough et al., 1965, Nielsen, 1999, 2000). Consistent with these intermediate levels of self-generated cognition and dreaming, NREM2 shows recruitment of numerous default network regions key to self-generated thought, including medial prefrontal cortex (BA 32), lateral temporal cortex bilaterally, and posterior cingulate cortex. Visual network regions also showed modest activation, particularly in the lingual gyrus (BA 18).
The frequency of dream experience is lowest in NREM3/4 (SWS) (Hobson et al., 2000, Nielsen, 2000). Indeed, significant numbers of participants can never recall cognitive activity or dreaming of any kind from SWS awakenings, despite multiple nights spent in the laboratory (Cavallero et al., 1992). Consistent with this modest level of self-generated thought, SWS showed some modest clusters of activation in various default network regions, including medial temporal lobe and left lateral temporal cortex, as well as small activations in visual network areas, including the middle occipital gyrus and lingual gyrus. Tempering these modest activations, however, were much larger and statistically stronger clusters of deactivation throughout major default network hubs, including medial prefrontal cortex (BA 32) and posterior cingulate cortex (BA 31). Further, deactivations were apparent in multiple subcortical brain areas, including the hypothalamus, thalamus, and dorsal pons. These subcortical deactivations are consistent with the overall decreased arousal and blockade of sensory inputs in SWS (Hobson et al., 2000).

Consistent with our predictions and prior work (Fox et al., 2013), REM, the sleep stage with by far the highest rates of dreaming (80-90% of the time (Hobson et al., 2000), showed heightened activation in numerous regions implicated in self-generated thought and imagery, especially widespread activation of the medial prefrontal cortex, the medial temporal lobe, and medial occipital areas. All of these activations are consistent with the self-generation of a self-referential narrative situated in a largely visual imaginal world.

These overall trends in activation patterns across the sleep stages are paralleled by changes in functional connectivity: connectivity among default mode network hubs, for instance, decreases monotonically throughout the NREM sleep stages (Sämann et al., 2011, Wilson et al., 2015). Consistent with these results, PET investigations have found a monotonic decrease in
cerebral energy metabolism across NREM stages 1-3 (Maquet, 1995), whereas energy metabolism in REM sleep is equal to (Maquet et al., 1990, Madsen et al., 1991, Braun et al., 1997) or higher than (Heiss et al., 1985, Buchsbaum et al., 1989) waking rest.

Finally, lucid REM sleep (LREM), in stark contrast to NREM sleep, shows many activations greater even than non-lucid REM sleep. The most striking difference is the reappearance of activity in the frontoparietal control network, discussed in more detail in the following section. Also apparent is heightened activity in areas already hyperactive in REM sleep, including medial prefrontal cortex and a large swathe of medial occipitoparietal cortex – potentially explained by anecdotal reports that lucid REM sleep experiences are much more vivid and detailed than regular REM dreams (Green, 1968, Sergio, 1988, Yuschak, 2006, Dresler et al., 2012).

**Brain deactivation, cognitive control, and meta-awareness throughout the sleep cycle**

As discussed above (Chapter 2), waking self-generated thought involves a co-activation of default network areas alongside executive frontoparietal control network regions, most notably dorsal anterior cingulate cortex, rostrolateral prefrontal cortex, and anterior inferior parietal lobule. These results are not particularly difficult to rationalize when it is recalled that cognitive control and meta-awareness, the principal functions tied to the latter network, are in fact quite prevalent in waking self-generated thought (even if lower than in typical externally-directed cognition and tasks; recall the discussion in Chapter 1 on cognitive control and metacognition in waking self-generated thought).

Conversely, executive and metacognitive functioning is largely absent or deficient in NREM and REM sleep cognition. Although dreams reports show strong thematic continuity with
the emotional and personal concerns of waking life (Cartwright et al., 1984, Kuiken et al., 2008, Fox et al., 2013), nonetheless actual goal-related thinking or top-down control and planning are rare. Further, activities involving sustained top-down control of attention, such as reading, writing, or using a phone or computer, occur only very rarely in dreams (Hartmann, 1996, Schredl, 2000). Although logical and paralogical thinking indeed take place in sleep and dreaming respects (Hall and Van de Castle, 1966, Kahn and Hobson, 2005), overall such thinking is only peripherally goal-related at best and is deficient in many other respects (Kahn and Hobson, 2005).

Metacognitive functioning is similarly compromised. Natural rates of meta-awareness of one’s true state during sleep (i.e., lucid dreaming or lucid sleep) are probably only about 1% even in experienced lucid dreamers (Snyder and Gackenbach, 1988, Schredl and Erlacher, 2004, 2011), and might occur only a handful of times throughout the entire lifespan in normal individuals (Barrett, 1991, Zadra et al., 1992). Moreover, employing a variety of interventions in an effort to increase meta-awareness during sleep (including psychological training, pharmacological agents, and external electrical stimulation), even among highly-motivated participants, tends to result in only very modest and poorly-validated gains (Stumbrys et al., 2012). Meta-awareness of other features of experience, such as bizarre or impossible situations and discontinuities of time and place (Dorus et al., 1971), is likewise heavily compromised (Kahn and Hobson, 2005).

Consistent with these many executive and metacognitive deficiencies throughout the sleep cycle, all sleep stages we investigated showed evidence for deactivation of frontoparietal control network regions. NREM1 showed deactivation of dorsolateral prefrontal cortex (BA 9) as well as premotor cortices (BA 6) bilaterally. In NREM2, dorsal anterior cingulate cortex (BA
24) and premotor cortex (BA 6) were similarly deactivated. SWS showed widespread deactivation of the executive system, including in orbitofrontal cortex bilaterally (BA 47), right rostrolateral prefrontal cortex (BA 10), dorsal anterior cingulate cortex (BA 32), and left ventrolateral prefrontal cortex (BA 47). Finally, REM sleep similarly showed large clusters of deactivation in the dorsolateral prefrontal cortex (BA 9/46) bilaterally, rostrolateral prefrontal cortex (BA 10) bilaterally, and ventrolateral prefrontal cortex (BA 47), also bilaterally. In stark contrast to all of these results throughout the rest of the sleep cycle, the unusual state of lucid REM (LREM) sleep instead showed activation of frontoparietal control network regions, including rostrolateral and dorsolateral prefrontal cortices, as well as in the anterior inferior parietal lobule bilaterally – consistent with restored cognitive control and meta-awareness (Dresler et al., 2012, Fox and Christoff, 2014).

Overall, these findings can be summarized in a tentative model of how subjective experience and sense of control/meta-awareness co-vary throughout the sleep cycle with activation in various brain regions that fall mostly with the default, visual, and frontoparietal control networks (Fig. 6). The coupling between default and visual network activation and the intensity of cognitive activity/dreaming on the one hand, and the orthogonal but equally intriguing coupling of frontoparietal control network activity with level of cognitive control and meta-awareness, is readily apparent (Fig. 6).
Figure 6. A speculative model of brain-mind isomorphisms across the sleep cycle. Brain activation and subjective experiences in waking self-generated thought (second point from left) are taken as an arbitrary baseline or zero-point, in relation to which brain activation and subjective experiences in other states of sleep and waking are compared. The \( y \)-axis employs arbitrary units for illustrative purposes only, to show relative brain activation/deactivation or increases/decreases in subjective intensity in each sleep stage, relative to waking self-generated thought. All values are therefore relative to an origin point defined by waking self-generated thought. Data are drawn from the meta-analyses detailed in the results section above, and from the review of content reports across sleep stages described in the introductory section. Some trends are apparente: for instance, DMN and VN activity broadly track subjective reports of the frequency and vividness of self-generated experience for each sleep stage: sleep stages with greater and more vivid subjective experiences and imagery tend to have greater activation in DMN and VN areas; conversely, SWS, in which subjective experiences and imagery are extremely rare, shows deactivations in DMN regions. A similar correspondence is shown between levels of cognitive control and meta-awareness and FPCN activation: FPCN activation is high, for instance, in externally-directed tasks where cognitive control is correspondingly high, but both cognitive control and FPCN activation fall in most sleep stages; however, some limited recovery of FPCN activation and cognitive control take place in so-called ‘lucid’ dreaming (far right point), in which the subject attains control and metacognitive awareness of the dream state. For further details, see the results and discussion section, above. DMN: default mode network; FPCN: frontoparietal control network; SWS: slow wave sleep; VN: visual network.

Limitations

Four major limitations of our results and speculative model should be emphasized. The first is that the subjective reports of nighttime dreaming and cognition have been collected
largely independently of the functional neuroimaging data that speaks to brain recruitment during the various sleep stages. That is, although differences in subjective experience across sleep stages are in general reliable and well-replicated (Nielsen, 1999, Hobson et al., 2000, Nielsen, 2000), the studies that have used functional neuroimaging to examine these same sleep stages have rarely actually collected dream or mentation reports from their participants (for instance, of the 8 studies we meta-analyzed for REM sleep, only one confirmed dreaming had indeed been taking place in the REM sleep periods examined in the PET scanner: (Maquet et al., 1996)). In the absence of such reports collected directly following functional neuroimaging of sleep and dreaming, any putative relationship between sleep stage neurophysiology and subjective content remains, at best, inferential and probabilistic. The obvious solution to this problem is for future functional neuroimaging studies of sleep to actively awaken and interrogate participants as to their subjective experiences across various sleep stages.

A second major concern is the small number of studies \((k)\) in our meta-analytic sample. Although we took pains to search the literature thoroughly and include every well-controlled and rigorously executed study, nonetheless the field of neuroimaging of sleep remains small. Accumulating research in this domain, however, will gradually mitigate this concern, as more powerful and representative syntheses become possible.

Third is the fact that most studies sleep deprived their participants the night before brain scanning in order to facilitate the maximum amount of sleep in the scanner. Aside from nonspecific effects, such as stress, sleep deprivation is known to affect the architecture and EEG correlates of the sleep cycle (Borbély et al., 1981), and might therefore influence neuroimaging measures of brain recruitment as well. In principle this concern could be addressed by comparing brain activation and deactivation for given sleep stages across studies that did and did not
employ sleep deprivation, but a much larger body of research is required before any such comparison is possible.

A fourth and final concern that should be reiterated is that sleep stages are as much a convenient abstraction as they are a concrete neurophysiological fact. We refer the reader back to the introductory material for a more detailed consideration of this important issue.

Conclusions

In a major review more than 15 years ago, Hobson et al. (2000) defended the position that REM and NREM sleep stages “can be defined, that their components can be analyzed and measured, and that they are significantly different from one another” (original emphasis). We have here provided substantial new evidence in favor of this view from a comprehensive meta-analysis of many functional neuroimaging studies across every stage of the sleep cycle. We have presented preliminary evidence that subjective experiences and neurophysiological markers covary across the sleep cycle, with the most intriguing finding being that DMN and visual network activation tracks the occurrence and immersiveness of self-generated thought, whereas frontoparietal control network deactivation tracks the general loss of cognitive control and meta-awareness of one’s psychological state. Overall, these results provide an intriguing example of complex, but nonetheless coherent, patterns of brain-mind isomorphism (Cacioppo and Tassinary, 1990). Our neuroimaging findings and accompanying theoretical account will hopefully serve as a useful stepping-stone on the path to a much deeper understanding of sleep and dream neurophysiology that surmounts the major limitations inherent in present research.
CHAPTER 4 – STUDY 3: SPECIFIC NEURAL RECRUITMENT ASSOCIATED WITH DIFFERENT TYPES OF THOUGHT

Introduction

Chapter 2 presented a detailed meta-analysis of general brain recruitment associated with self-generated forms of thought, but could offer only speculations about the specific functional roles played by the various regions and networks consistently recruited. Major questions remain, however, about the specific neural recruitment associated with specific types of thought – questions that cannot be addressed, beyond simple reverse inference, without targeted experiments.

Investigating abstract qualities of thought

In the Introduction (Chapter 1), we reviewed the bewildering variety of first-person content that characterizes self-generated thought. This multiplicity of specific thought content can be simplified to some extent by examining higher-level, more abstract qualities of self-generated thought. Rather than referring to specific content, abstract qualities of thought involve the high-level functionality or category to which the thought can be ascribed. Some major examples of abstract qualities of thought under which virtually an infinite variety of specific contents can be categorized are summarized in Table 12. Aside from assessing the relative frequency of each of these thought qualities in a healthy adult sample, another aim was to assess their co-dependence (if any). In principle each of these major qualities should be orthogonal to the others – goal-related thinking should be able to arise spontaneously or be intentionally directed, for instance, and ought to be able to feel either pleasant or unpleasant, depending on the
specific thought content. Nonetheless, there remains the possibility that thought qualities might co-vary.

The first major quality we sought to investigate is the relation of thought to the immediate environment, or its stimulus-dependence. Here, we use stimulus in the general sense outline in the Introduction (Chapter 1), to encompass any sensory input from the external environment, with external operationalized in its literal sense as beyond the boundary of the body. Stimulus-dependent thought, then, is any thought directly related to or triggered by incoming visual, auditory, tactile, olfactory, or gustatory sensation. Many studies, including the present investigation (see Methods, below), deliberately minimize sources of such external input in order to increase the relative frequency of stimulus-independent thought – the object of interest. Stimulus-independent thought, then, is any other form of cognitive-affective activity not directly related to or triggered by the surrounding ‘here and now.’ Our aim in this study was to discard trials in which subjects reported experience stimulus-dependent thought about their environments, and instead to resolve different types of stimulus-independent thought with follow-up questions in the fMRI scanner.

The first major quality of stimulus-independent thought we investigated was agency: i.e., the subjective assessment of whether one’s thought is being intentionally and volitionally directed/guided, or arises and unfolds in a spontaneous fashion. Note that here we use the term ‘spontaneous’ in the limited sense of not being consciously directed or controlled – elsewhere we have developed a technical definition of spontaneous thought that is beyond the scope of the present investigation (Christoff et al., in press).

A third central quality is the self-assessed utility of thought, i.e. whether thought is judged to be related to one’s current goals and concerns or not. The rationale for addressing this
question draws on a large body of psychological research dating back to the early days of mind-wandering research, which collectively suggests that goal- and concern-related thinking plays a large, perhaps dominant, role in self-generated thought (Klinger, 1987, Klinger and Cox, 2004, 2011, McMillan et al., 2013). The prevalence of goal- and concern-related thinking, present in upwards of half of all thoughts, has also spawned a conceptual model of mind-wandering and related forms of thinking, the Current Concerns Theory, which posits that such goal-related thinking is a central feature, and adaptive quality, of self-generated thought (Klinger, 1987, Klinger and Cox, 2004, 2011, McMillan et al., 2013).

Finally, we investigated the valence of self-generated thought (from unpleasant, to neutral, to pleasant), in line with extensive research showing that a wide variety of affect accompanies self-generated thought of various kinds (Fox et al., 2014, Fox et al., in preparation).

**Table 12.** Major qualities of self-generated thought uninvestigated at the neural level.

<table>
<thead>
<tr>
<th>Quality</th>
<th>Major Thought Alternatives</th>
<th>Overview</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stimulus-dependence</strong></td>
<td>Stimulus-independent</td>
<td>Thoughts not caused by, or not directly related to, the surrounding environment</td>
</tr>
<tr>
<td></td>
<td>Stimulus-dependent</td>
<td>Thoughts caused by, or directly related to, the surrounding environment</td>
</tr>
<tr>
<td><strong>Agency</strong></td>
<td>Spontaneously Unfolding</td>
<td>Thoughts unfolding without one’s direct control or guidance</td>
</tr>
<tr>
<td></td>
<td>Intentionally Directed</td>
<td>Thoughts unfolding with one’s direct control or guidance</td>
</tr>
<tr>
<td><strong>Utility</strong></td>
<td>Goal-related</td>
<td>Thoughts related in some way to one’s goals or current concerns</td>
</tr>
<tr>
<td></td>
<td>Goal-unrelated</td>
<td>Thoughts unrelated one’s goals or current concerns</td>
</tr>
<tr>
<td><strong>Valence</strong></td>
<td>Negative</td>
<td>Unpleasant, negatively-valenced thoughts</td>
</tr>
<tr>
<td></td>
<td>Neutral</td>
<td>Thoughts without an obvious affective component or flat affect</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>Pleasant, positively-valenced thoughts</td>
</tr>
</tbody>
</table>
Previous efforts to understand specific neural recruitment associated with different types of thought

Only a very limited number of studies have addressed brain recruitment associated with specific thought content or abstract qualities, for reasons that will become apparent in the Methods section. Briefly, the central problem is the low signal-to-noise ratio of the fMRI BOLD signal, which necessitates a large number of trials in order for event-related designs to yield robust results. Coupled with the inability to experimentally control the frequency, much less the specific content, of thought, this produces a situation in which it is very difficult to collect an adequate number of trials for multiple types of thought within a reasonable scanning time.

Some preliminary work along these lines has been reported, however. A recent study by Vanhaudenhuyse and colleagues (2011) investigated the involvement of three key default network regions in the ‘intensity’ of internal (vs. external, i.e. perceptual) awareness. They defined internal awareness as “all environmental stimuli-independent thoughts (e.g., inner speech, autobiographical memories, or wandering thoughts).” During continuous fMRI acquisition, subjects were prompted by quasi-random experience-sampling probes that asked them to rate the contents of their awareness on a 4-point scale from strongly external (‘1’) or moderately external (‘2’) through to moderately internal (‘3’) and strongly internal (‘4’). Investigating three key default network regions-of-interest (the medial prefrontal cortex, posterior cingulate cortex, and bilateral medial temporal lobe), they found that activity in these regions correlated positively with the self-reported intensity of internal awareness (Vanhaudenhuyse et al., 2011). This study presents some intriguing preliminary evidence that specific neural correlates accompany thought that is specifically stimulus-independent, but the region-of-interest approach limits its generalizability and reliability.

Another study has investigated different types of affect in self-generated thought. Tusche
and colleagues (2014) investigated the emotional content of unconstrained thoughts during task-free rest periods using multivariate pattern analysis. They found that patterns of activity within medial prefrontal cortex predicted the valence (positive vs. negative) of thoughts – not just during the initial session, but also at a follow-up scanning session one week later (Tusche et al., 2014).

To our knowledge, no study has directly investigated neural correlates of intentionally directed/guided thought, or of goal- and concern-related thoughts, and the aforementioned studies investigating stimulus-independence and valence clearly leave many questions unanswered. The present study therefore aimed to improve our understanding of self-generated thought by resolving three major categories of stimulus-independent thought at the neural level.

**Methods**

**Participants**

Participants were 16 healthy young adults (9 female; mean age ± SD in years = 28.1 ± 4.3). Most participants were graduate students at the University of British Columbia, recruited via snowball (i.e., word-of-mouth referral) sampling. All participants selected for participation in the study were neurologically and psychiatrically healthy; right-handed (to minimize any functional or structural neuroanatomical differences due to language lateralization); and provided written informed consent in accordance with the ethical standards set by the Clinical Research Ethics Board (CREB) of the University of British Columbia, Vancouver, Canada. The study protocol was reviewed and approved by the UBC MRI Research Centre. Participants were compensated $20/hr for their time.
Experimental design

As discussed at length above (Chapter 2), general neural correlates of self-generated thought have been identified in some two dozen studies, and in at least ten agnostic, whole-brain analyses. These investigations have converged on about a dozen regions that appear to be consistently recruited (Fox et al., 2015), but almost nothing is known about the functional roles of each of these regions except via inference from other investigations that have interrogated a given region’s functionality using task-based designs (Andrews-Hanna et al., 2014). The main hurdle to investigating multiple types of self-generated thought with fMRI is in obtaining adequate statistical power: well-powered event-related fMRI analyses require ~15-20 trials for a given event type, with >25 trials being ideal (Murphy and Garavan, 2005). The problem is that self-generated thought, by definition, cannot be controlled or manipulated by the experimenter; a sufficient number of probes, and consequently time in the scanner, must be allowed to elapse such that a sufficient quantity of each thought type arises spontaneously (Vanhaudenhuyse et al., 2011). Our solution to this problem was to allow for the maximal number of thought probes by (i) repeating thought probes as quickly as possible, provided that they continued to yield high rates of self-reported thought, and did not result in frustration or overstimulation for participants; and (ii) dispensing with any task that would sacrifice scanning time.

In order to (i) maximize the frequency of thought probes without participants feeling overwhelmed, we conducted a pilot behavioral study in which inter-probe intervals (IPIs) of varying length (4, 6, 8, 10, 12, 14, 16, 18, and 20 s) were randomly utilized throughout 5 blocks of 24 probes each. Although rates of self-reported thought were very high (>90%) under all interval lengths, participants nonetheless reported that the 4 and 6 s intervals were too brief and led to confusion and an inability to become immersed in the stream of thought. We therefore
used a randomized range of IPIs (8, 10, 12, 14, 16, and 18 s; mean = 13 s), with equal numbers of each IPI, for each block and each participant.

In order to further maximize the scanning time dedicated to self-generated thought of various kinds, we (ii) dispensed with a task altogether. Most prior studies of self-generated thought have employed some simple and boring task paradigm, with the main aims being to induce mind-wandering or other task-unrelated thought, and additionally to provide some objective corroboration of mind-wandering episodes via observable decrements in accuracy or other measures of performance (e.g., reaction time).

As noted in many contemplative traditions, however, thought in no way needs to be ‘induced’ – the unoccupied mind will naturally produce a nearly nonstop stream of thoughts, images, inner speech, and so on (Gunaratana, 2011, Eifring, 2015). Moreover, ‘objective’ measures of attentional lapses do not correlate particularly well (or, often, at all) with actual subjective reports of mind-wandering or being off-task; especially with highly practiced tasks, self-generated thought can co-exist perfectly well with adequate task performance. For instance, in a recent fMRI investigation of mind-wandering from our research group, no significant differences whatsoever were observed in either reaction time or accuracy on a simple sustained attention task when ‘on-task’ vs. ‘off-task’ (i.e., mind-wandering) epochs were compared (Christoff et al., 2009). Moreover, even if a suitable objective behavioral or physiological measure of self-generated thought were known, it would likely reveal nothing about the specific content of thought, which was the object of our investigation here. A task was therefore dispensed with altogether in the present study.

In summary, to date no fMRI study has adequately examined multiple types of self-generated thought, because the number of trials required is very large and is uncontrollable when
allowing the subject to self-generate thought content; therefore our design allowed for the maximal number of trials (120) within a reasonable scanning timeframe (1.5 hrs, including acquisition of 3D T1 and DTI white matter images for morphometric neuroimaging analysis; see Chapter 6) so that individual thought types could be resolved with sufficient statistical power in the ensuing analyses.

Each trial (Fig. 7) involved the participant lying quietly, fixating on a blank black screen for between 8–18s (the ‘thought period’). Thought periods were then randomly interrupted with thought probes. The first probe asked participants whether, just prior to the probe onset, they had been having stimulus-independent, stimulus-dependent, or no thoughts. If the participant responded to either of the latter two options, the trial was terminated and a new ‘thought period’ began. If the participant indicated stimulus-independent thoughts, however, they then proceeded to answer three additional probes about the qualities of these thoughts.

**Procedure and fMRI thought sampling paradigm**

First, all participants reported to our laboratory space at the Brain Research Centre (BRC) inside the UBC Hospital. Upon arrival at the lab space, participants completed a consent form, and were then given a detailed explanation of the study and the thought probe questions by the experimenter. A scripted explanation was used to provide identical examples of each thought type for every participant. Participants were then left alone to engage in two practice blocks (24 probes each) of the study (~15 min), during which they had an unlimited time to classify their thoughts and respond to each probe (Fig. 7). E-Prime 2 was used to present the thought probes (Psychology Software Tools Inc., Sharpsburg, PA). If the participant responded to the first probe that they could recall ‘No thoughts’ or that their thoughts were ‘Stimulus-dependent,’ then that
trial ended and a new period of a blank black screen (‘thought period’) began. As noted above, our reason for ignoring stimulus-dependent thoughts was to ensure the maximal amount of scanning time dedicated to stimulus-independent thoughts, instead.

Following the practice blocks, the experimenter discussed the experiment and thought probe questions further with each participant in order to clarify any ambiguities in the questions or in the way in which the individual’s thoughts should be classified. Because each probe had a limited number of discrete responses (indeed, response options were mostly binarized), it was emphasized to participants that they should in every case choose the best answer, acknowledging that in some cases thoughts might fall somewhere in between the two response options provided and might not be easily classifiable.

When all questions had been clarified, the participant was escorted to the UBC MRI Research Centre, cleared through screening and security protocols, and entered the MRI with the aid of a staff technician. The exact same thought sampling task as given in the practice blocks was projected onto a screen within the MRI (Fig. 7), and responses were made using an MRI-compatible button-box instead of a standard computer keyboard (as in the practice session). The thought sampling experiment was divided into 5 blocks of 24 probes each, for a total of 120 probes. Each block proceeded automatically, with 8 s allowed for each response to each probe question. At the end of each block the experimenter checked in with each participant as to their physical and mental comfort in the scanner, and allowed participants to take a brief break. Upon completion of the experiment, participants were debriefed, and were compensated $20/hr for their time. Most participants returned within 1-2 days to the BRC lab space to complete an identical behavioral version of the thought sampling paradigm for correlational use with our anatomical scans (for details, see Chapter 6).
Figure 7. Thought probe questions and sequence (identical for both the behavioral and fMRI versions of the experiment).

**fMRI data acquisition**

The order of acquisition was first a 3D T1-weighted anatomical scan (~6 min), followed by five T2*-weighted functional runs (~8 min each), and ending with a diffusion tensor imaging (DTI) scan (~8 min). Functional and structural MRI data were collected using a 3.0 Tesla Philips Intera MRI scanner (Best, Netherlands) with a standard head coil. Head movement was restricted using foam padding around the head. T2*-weighted functional images were acquired parallel to the anterior commissure/posterior commissure (AC/PC) line using a single-shot gradient echo-planar sequence (EPI; repetition time \(TR\) = 2 s, echo time \(TE\) = 30 ms, flip angle \(FA\) = 90°, field of view \(FOV\) = 240 × 240 × 143 mm, matrix size = 80 × 80, SENSE factor = 1.0). The exact number of functional volumes varied for each run for each subject, since the pacing of the thought probe responses was self-timed, but was approximately 6 to 8 minutes (i.e., between 180 and 240 functional volumes per run). A longer scan (12 min) was initiated for each run, so that the participant would not run out of time and have the scan end prematurely; when it was observed that the participant had completed the run in the behavioral software (E-Prime), the scan was prematurely force-terminated. The resulting truncated functional volumes were manually inspected and edited to remove partial volume acquisitions. Each functional volume
included 36 interleaved axial slices (3 mm thick with 1 mm skip) covering the entire brain.

Before functional imaging, a 3D T1-weighted structural volume was acquired, and following the functional scans, a DTI image was acquired. Details of these scans are reported in a later chapter detailing our morphometric neuroimaging experiments (Chapter 6).

*fMRI data preprocessing and analysis*

fMRI data for each subject were preprocessed and analyzed using SPM8 (Statistical Parametric Mapping, Wellcome Department of Imaging Neuroscience, London, UK). Slice timing correction was performed using sinc interpolation and resampling with the middle (18th) slice as a reference point. All functional volumes were realigned to the first volume to correct for between-scan motion. The structural volume was coregistered to the mean functional image and segmented to extract a gray matter image. The segmented structural volume was then spatially normalized to a gray matter image of the Montreal Neurological Institute (MNI) template and resliced to a voxel size of $2 \times 2 \times 2$ mm. The derived spatial transformations were applied to the 20 realigned functional volumes to bring them into standardized MNI space. Finally, the functional volumes were smoothed with an 8-mm full-width at half-maximum (FWHM) isotropic Gaussian kernel to compensate for residual between-subject variability after spatial normalization and to permit application of Gaussian random field theory for corrected statistical inference (Friston et al., 1994). To remove low-frequency drift in the blood oxygen-level dependent (BOLD) signal, the data were high-pass filtered using an upper cut-off period of 128 s. No global scaling was performed.

Condition effects at each voxel were estimated according to the general linear model in our whole-brain analyses. The model included (i) the observed time series of intensity values at
each voxel, which represented the dependent variable; (ii) covariates modeling session-specific 
effects (i.e., the six head movement parameters), treated as variables of no interest in our 
contrasts; and (iii) regressor functions constructed by convolving thought-type-specific stick 
functions with a synthetic hemodynamic response function. The regressor functions were 
constructed to model each of the different types of reported thought for each block (i.e., 
stimulus-independent, stimulus-dependent, no thoughts, spontaneously unfolding, intentionally 
directed, goal-related, goal-unrelated, and unpleasant, neutral, and pleasant). Valence (affect) 
was presented as a five-point scale in the actual thought sampling paradigm, but all data was 
collapsed into three categories (negative, neutral, and positive) to increase power. The onset of 
the thoughts was set at 8 s prior to the onset of the first thought probe, and modeled for a 
duration of 8 s. The onset and responses to the thought probes themselves were also modeled and 
treated as regressors of no interest. As discussed above, data relating to stimulus-dependent 
thoughts was of no interest in the present study and was therefore also modeled as a regressor of 
no interest.

Group random-effects analyses were then performed for each contrast to compare 
different types of stimulus-independent thought to one another (i.e., spontaneously unfolding vs. 
intentionally directed thoughts, goal-related vs. goal-unrelated thoughts, negative vs. neutral + 
positive thoughts, and positive vs. neutral + negative thoughts). Due to the relatively exploratory 
nature of the study, the threshold for significance was set at $p < 0.001$ (uncorrected for multiple 
comparisons) and extent threshold $k > 5$ voxels; for visualization purposes in our figures, this 
threshold was relaxed to $p < 0.005$ (uncorrected) and $k > 5$ (Christoff et al., 2009).
Results

Behavioral results

Behavioral data for all participants, collapsed across all 5 fMRI runs, is presented in Table 13. As our interest in this study was in the neural correlates particular thought types, rather than individual differences in thought tendencies, this data is not analyzed further here. For more discussion and analysis, however, see Chapters 5 and 6.

Table 13. Behavioral data for all fMRI runs.

<table>
<thead>
<tr>
<th>Subject</th>
<th>% SIT</th>
<th>% Spontaneous</th>
<th>% Related</th>
<th>% (--)</th>
<th>% Neutral</th>
<th>% (+)</th>
<th>Age</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>S03</td>
<td>81.74</td>
<td>82.98</td>
<td>68.09</td>
<td>30.85</td>
<td>23.40</td>
<td>45.74</td>
<td>22</td>
<td>M</td>
</tr>
<tr>
<td>S05</td>
<td>80.91</td>
<td>80.23</td>
<td>49.40</td>
<td>4.55</td>
<td>65.91</td>
<td>29.55</td>
<td>25</td>
<td>M</td>
</tr>
<tr>
<td>S06</td>
<td>84.76</td>
<td>70.93</td>
<td>28.57</td>
<td>17.24</td>
<td>52.87</td>
<td>29.89</td>
<td>24</td>
<td>F</td>
</tr>
<tr>
<td>S07</td>
<td>95.00</td>
<td>81.58</td>
<td>24.11</td>
<td>30.09</td>
<td>33.63</td>
<td>36.28</td>
<td>34</td>
<td>F</td>
</tr>
<tr>
<td>S08</td>
<td>84.87</td>
<td>96.04</td>
<td>26.73</td>
<td>19.00</td>
<td>52.00</td>
<td>29.00</td>
<td>35</td>
<td>F</td>
</tr>
<tr>
<td>S09</td>
<td>71.93</td>
<td>100.00</td>
<td>36.25</td>
<td>3.75</td>
<td>65.00</td>
<td>31.25</td>
<td>28</td>
<td>M</td>
</tr>
<tr>
<td>S10</td>
<td>79.66</td>
<td>76.60</td>
<td>65.59</td>
<td>1.09</td>
<td>43.48</td>
<td>55.43</td>
<td>33</td>
<td>M</td>
</tr>
<tr>
<td>S11</td>
<td>85.34</td>
<td>50.51</td>
<td>33.33</td>
<td>12.12</td>
<td>51.52</td>
<td>36.36</td>
<td>28</td>
<td>F</td>
</tr>
<tr>
<td>S12</td>
<td>82.76</td>
<td>68.09</td>
<td>57.45</td>
<td>46.24</td>
<td>33.33</td>
<td>20.43</td>
<td>35</td>
<td>F</td>
</tr>
<tr>
<td>S13</td>
<td>71.29</td>
<td>87.32</td>
<td>8.70</td>
<td>15.49</td>
<td>70.42</td>
<td>14.08</td>
<td>24</td>
<td>F</td>
</tr>
<tr>
<td>S14</td>
<td>89.17</td>
<td>60.75</td>
<td>47.66</td>
<td>11.21</td>
<td>72.90</td>
<td>15.89</td>
<td>26</td>
<td>F</td>
</tr>
<tr>
<td>S15</td>
<td>88.98</td>
<td>100.00</td>
<td>55.24</td>
<td>12.38</td>
<td>33.33</td>
<td>54.29</td>
<td>26</td>
<td>F</td>
</tr>
<tr>
<td>S17</td>
<td>95.83</td>
<td>98.26</td>
<td>81.74</td>
<td>26.09</td>
<td>13.91</td>
<td>60.00</td>
<td>26</td>
<td>F</td>
</tr>
<tr>
<td>S18</td>
<td>82.73</td>
<td>60.23</td>
<td>52.75</td>
<td>13.19</td>
<td>32.97</td>
<td>53.85</td>
<td>28</td>
<td>M</td>
</tr>
<tr>
<td>S19</td>
<td>77.53</td>
<td>66.18</td>
<td>30.43</td>
<td>5.97</td>
<td>65.67</td>
<td>28.36</td>
<td>31</td>
<td>M</td>
</tr>
<tr>
<td>S31</td>
<td>94.17</td>
<td>62.50</td>
<td>36.28</td>
<td>23.64</td>
<td>26.36</td>
<td>50.00</td>
<td>23</td>
<td>M</td>
</tr>
<tr>
<td>Mean</td>
<td>84.20</td>
<td>77.64</td>
<td>43.89</td>
<td>17.06</td>
<td>46.04</td>
<td>36.90</td>
<td>28</td>
<td>–</td>
</tr>
</tbody>
</table>

SIT: stimulus-independent thought.

Orthogonality of thought qualities

To examine whether abstract thought qualities were independent of one another or not, we computed all pairwise correlations, for all thought responses, for all participants (across 120 probe responses). This resulted in three correlation coefficients per subject, for relationships between spontaneity/goal-relatedness; spontaneity/affect; and goal-relatedness/affect. All coefficients were Fisher-transformed (Fisher, 1915) and then averaged to determine the mean
relationship between each thought type. As summarized in Table 14, relationships between thought types were generally small and all correlations were non-significant (\(df = 15\), all \(p\)’s > .05). With 1-2 outliers (with \(r\)’s > .5) removed for each pairwise correlation, these relationships became even smaller, with no more than ~6% of the variance in one thought type explained by any other. Importantly, the weakest relationship was that between directedness/spontaneity of thought and goal-relatedness of thought. This was a critical finding because we anticipated that both intentional direction and goal-relatedness of thought might recruit executive resources, so establishing the orthogonality of these thought types was important to the validity of our subsequent fMRI contrasts.

**Table 14.** Pairwise correlations between thought qualities.

<table>
<thead>
<tr>
<th>Correlation Coefficient</th>
<th>Spontaneity-Relatedness</th>
<th>Spontaneity-Affect</th>
<th>Relatedness-Affect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (r)</td>
<td>-.180</td>
<td>.200</td>
<td>.273</td>
</tr>
<tr>
<td>Mean (r) (with outliers &gt;.5 removed)</td>
<td>-.118</td>
<td>.179</td>
<td>.249</td>
</tr>
<tr>
<td>(R^2) (with outliers &gt;.5 removed)</td>
<td>.014 (1.4% of variance explained)</td>
<td>.032 (3.2% of variance explained)</td>
<td>.062 (6.2% of variance explained)</td>
</tr>
</tbody>
</table>

**Neural correlates of spontaneity and directedness in stimulus-independent thought**

Our first contrast involved examining the sense of agency accompanying stimulus-independent thoughts, i.e., whether the thoughts were judged to be *spontaneously unfolding* or *intentionally directed*. Thoughts were judged to arise spontaneously (77.6% of the time) much more often than they were deliberately directed (22.4% of the time). The brain contrast for spontaneously unfolding thoughts (> intentionally directed) revealed an array of posterior default network regions that were more highly activated when thoughts unfolded spontaneously (Table
15, Fig. 8), including the temporoparietal junction bilaterally, the right parahippocampal cortex, right superior temporal gyrus, and precuneus/posterior cingulate cortex. The right putamen was also significantly activated.

No significant activations were observed for the contrast intentionally directed thoughts > spontaneously unfolding thoughts at the same threshold. At a slightly relaxed threshold (as used for visualization purposes in our figures: \( t > 2.9467, p < .005 \), uncorrected, \( k > 5 \)), a single cluster of activation was observed, localized to the white matter adjacent to the left dorsal anterior cingulate cortex (\( t = 4.10, Z = 3.31, k = 12 \), MNI coordinates: -24, 11, 28; Fig. 9).
### Table 15. Activation peaks for spontaneously unfolding thoughts (> intentionally directed thoughts)

<table>
<thead>
<tr>
<th>Region</th>
<th>Side</th>
<th>BA</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxels</th>
<th>$t$-statistic</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temporoparietal junction</td>
<td>R</td>
<td>40/41/42/13</td>
<td>57</td>
<td>-40</td>
<td>19</td>
<td>79</td>
<td>5.41</td>
<td>3.97</td>
</tr>
<tr>
<td>Temporoparietal junction/posterior insula</td>
<td>L</td>
<td>13/41/22</td>
<td>-45</td>
<td>-28</td>
<td>7</td>
<td>88</td>
<td>5.31</td>
<td>3.92</td>
</tr>
<tr>
<td>Putamen</td>
<td>R</td>
<td>–</td>
<td>24</td>
<td>14</td>
<td>-2</td>
<td>9</td>
<td>4.84</td>
<td>3.70</td>
</tr>
<tr>
<td>Parahippocampal cortex</td>
<td>R</td>
<td>35/36</td>
<td>27</td>
<td>-31</td>
<td>-20</td>
<td>14</td>
<td>4.68</td>
<td>3.62</td>
</tr>
<tr>
<td>Superior temporal gyrus</td>
<td>R</td>
<td>22/21/38</td>
<td>54</td>
<td>-10</td>
<td>-8</td>
<td>6</td>
<td>4.07</td>
<td>3.29</td>
</tr>
<tr>
<td>Precuneus/posterior cingulate cortex</td>
<td>M</td>
<td>7/31</td>
<td>12</td>
<td>-49</td>
<td>49</td>
<td>14</td>
<td>4.04</td>
<td>3.27</td>
</tr>
</tbody>
</table>

Height threshold $t = 3.73$ ($p < .001$, uncorrected), extent threshold $k > 5$. 

---

88
Figure 8. Activations for spontaneously unfolding thoughts > intentionally directed thoughts. IPL: inferior parietal lobule; Ins: insula; LG: lingual gyrus; LTC: lateral temporal cortex; PCC: posterior cingulate cortex; Prec: precuneus; Put: putamen; STG: superior temporal gyrus; TPJ: temporoparietal junction. Height threshold $t = 2.95$, $p < .005$, uncorrected; extent threshold $k > 5$. 
Figure 9. Sub-threshold activation in white matter adjacent to the dorsal ACC/anterior MCC for intentionally directed > spontaneously unfolding thoughts. Height threshold $t = 2.95$, $p < .005$, uncorrected; extent threshold $k > 5$. 
Neural correlates of goal-relatedness in stimulus-independent thought

Twelve significant clusters of activation were observed for the contrast goal-related thoughts > goal-unrelated thoughts (Table 16; Fig. 10).

Table 16. Activation peaks for goal-related thoughts (> goal-unrelated thoughts)

<table>
<thead>
<tr>
<th>Region</th>
<th>Side</th>
<th>BA</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxels</th>
<th>t-statistic</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>White matter adjacent to cingulate cortex</td>
<td>R</td>
<td>27</td>
<td>-10</td>
<td>31</td>
<td>83</td>
<td>6.65</td>
<td>4.47</td>
<td></td>
</tr>
<tr>
<td>Cerebellum</td>
<td>R</td>
<td>15</td>
<td>-40</td>
<td>-41</td>
<td>40</td>
<td>5.88</td>
<td>4.17</td>
<td></td>
</tr>
<tr>
<td>Lingual gyrus/fusiform gyrus</td>
<td>R</td>
<td>30</td>
<td>-70</td>
<td>-5</td>
<td>83</td>
<td>4.24</td>
<td>3.39</td>
<td></td>
</tr>
<tr>
<td>Thalamus</td>
<td>R</td>
<td>18</td>
<td>-16</td>
<td>1</td>
<td>21</td>
<td>5.52</td>
<td>4.02</td>
<td></td>
</tr>
<tr>
<td>Premotor cortex/Mid-cingulate cortex</td>
<td>L</td>
<td>6/24/32</td>
<td>-12</td>
<td>8</td>
<td>49</td>
<td>5.25</td>
<td>3.89</td>
<td></td>
</tr>
<tr>
<td>Lingual gyrus/fusiform gyrus</td>
<td>L</td>
<td>18/19</td>
<td>-21</td>
<td>-64</td>
<td>-11</td>
<td>46</td>
<td>5.00</td>
<td>3.78</td>
</tr>
<tr>
<td>Mid-cingulate cortex</td>
<td>L</td>
<td>31/24</td>
<td>-21</td>
<td>-22</td>
<td>40</td>
<td>21</td>
<td>4.67</td>
<td>3.61</td>
</tr>
<tr>
<td>Parahippocampal cortex/Cerebellum</td>
<td>R</td>
<td>36</td>
<td>30</td>
<td>-34</td>
<td>-26</td>
<td>7</td>
<td>4.64</td>
<td>3.60</td>
</tr>
<tr>
<td>Cerebellum (lobule VI)</td>
<td>L</td>
<td>29/26</td>
<td>-9</td>
<td>-37</td>
<td>13</td>
<td>7</td>
<td>4.14</td>
<td>3.33</td>
</tr>
<tr>
<td>Posterior cingulate cortex</td>
<td>M</td>
<td>23</td>
<td>9</td>
<td>-58</td>
<td>19</td>
<td>22</td>
<td>4.45</td>
<td>3.50</td>
</tr>
<tr>
<td>Retrosplenial cortex</td>
<td>M</td>
<td>18/19</td>
<td>-9</td>
<td>-64</td>
<td>7</td>
<td>7</td>
<td>3.93</td>
<td>3.21</td>
</tr>
</tbody>
</table>

Height threshold \( t = 3.73 \) (\( p < .001 \), uncorrected), extent threshold \( k > 5 \).

No significant activations were observed for the contrast goal-unrelated thoughts > goal-related thoughts.
Figure 10. Activations for goal-related thoughts > goal-unrelated thoughts. aMCC: anterior midcingulate cortex; dACC: dorsal anterior cingulate cortex; MCC: mid-cingulate cortex; mOcc: medial occipital cortex; PCC: posterior cingulate cortex; PMC: premotor cortex; Thal: thalamus; VI: cerebellar lobule 6; VIII: cerebellar lobule 8; WM: white matter. Height threshold $t = 2.95$, $p < .005$, uncorrected; extent threshold $k > 5$. 
Neural correlates of affective valence in stimulus-independent thought

Several clusters of activation were observed in the dorsal striatum for the contrast positive > (neutral + negative thoughts), particularly throughout the caudate nucleus. Activations were also observed in the right cerebellum and in the white matter adjacent to the left anterior cingulate cortex, as well as at more liberal thresholds in the orbitofrontal cortex/subgenual anterior cingulate cortex and precuneus (Table 17; Fig. 11).

Table 17. Activation peaks for positively-valenced thoughts (> neutral + negative thoughts)

<table>
<thead>
<tr>
<th>Region</th>
<th>Side</th>
<th>BA</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxels</th>
<th>t-statistic</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Caudate (head)</td>
<td>L</td>
<td></td>
<td>-18</td>
<td>29</td>
<td>-8</td>
<td>10</td>
<td>5.54</td>
<td>4.03</td>
</tr>
<tr>
<td>Caudate (tail)</td>
<td>L</td>
<td></td>
<td>-21</td>
<td>-37</td>
<td>13</td>
<td>22</td>
<td>5.02</td>
<td>3.79</td>
</tr>
<tr>
<td>Cerebellum (region)</td>
<td>R</td>
<td></td>
<td>30</td>
<td>-73</td>
<td>-47</td>
<td>34</td>
<td>4.71</td>
<td>3.63</td>
</tr>
<tr>
<td>White matter adjacent to anterior cingulate cortex</td>
<td>L</td>
<td>32/24</td>
<td>-18</td>
<td>20</td>
<td>28</td>
<td>11</td>
<td>4.49</td>
<td>3.52</td>
</tr>
<tr>
<td>Caudate (body)</td>
<td>L</td>
<td></td>
<td>-18</td>
<td>-13</td>
<td>19</td>
<td>6</td>
<td>4.20</td>
<td>3.36</td>
</tr>
</tbody>
</table>

Height threshold $t = 3.73$ ($p < .001$, uncorrected), extent threshold $k > 5$. 
Figure 11. Activations for positively-valenced thoughts > neutral and negatively-valenced thoughts. Caud: caudate nucleus; dACC: dorsal anterior cingulate cortex; HPC: hippocampus; PHC: parahippocampal cortex; Prec: precuneus; Thal: thalamus; WM: white matter. Height threshold $t = 2.95$, $p < .005$, uncorrected; extent threshold $k > 5$.

No significant activations were observed for the contrast negative thoughts > neutral + positive thoughts or for the contrast neutral thoughts > positive + negative thoughts.
Discussion

The preceding results, though preliminary, are the first to resolve several interesting categorical divisions of stimulus-independent and self-generated thought at the neural level. Specifically, we found that the qualities of agency, utility, and valence of stimulus-independent thoughts could all be resolved at the neural level.

**Neural correlates of spontaneously unfolding and intentionally directed thoughts**

For spontaneously unfolding thoughts, we found greater recruitment in a number of posterior default network regions, consistent with the notion that self-generated thoughts can arise without top-down control and capture attention in a bottom-up manner (Ellamil et al., 2016). The most intriguing result was a large cluster of activation in the right medial temporal lobe (parahippocampal cortex), which seems to be critically involved in self-generating thought: many intracranial EEG experiments in humans have demonstrated that stimulation of medial temporal lobe elicits thoughts, memories, and dream-like experiences (see Chapter 8), and activity in medial temporal lobe has been shown to peak just prior to the subjective experience of a spontaneously arising thought or memory, in both intracranial EEG (Gelbard-Sagiv et al., 2008) and fMRI (Ellamil et al., 2016) experiments. Our other results were highly consistent with a recent study we conducted specifically examining the spontaneous arising of thoughts (but not their intentional direction or guidance): similar to this investigation (Ellamil et al., 2016), we also found activations in the right inferior parietal lobule/temporoparietal junction, the posterior insula (albeit in the contralateral hemisphere), and posterior cingulate cortex. Inferior parietal lobule has been repeatedly implicated in bottom-up attentional capture by retrieved memories (Cabeza et al., 2008). The lateral temporal cortex and temporoparietal junction have been
hypothesized to be part of a particular default network subsystem involved in mentalizing and conceptual processing during self-generated thought (Andrews-Hanna et al., 2014), potentially therefore supporting spontaneous theory of mind and semantic memory retrieval (Kvavilashvili and Mandler, 2004).

In addition (and in contrast to the study by Ellamil and colleagues), we observed clusters in the lingual gyrus and superior temporal gyrus. Findings in ‘sensory’ cortices such as lingual gyrus, superior temporal gyrus, and posterior insula could be explained, respectively, by visual, auditory, and interoceptive thought and imagery in self-generated thought capturing attention in a bottom-up fashion (see Chapter 1 for a review of thought content detailing the high frequency of all these content types).

For intentionally directed thoughts, we observed only one subthreshold cluster, in the dorsal anterior cingulate cortex/anterior mid-cingulate cortex, a canonical executive control region. The meager results for this analysis are likely due to the small number of intentionally directed thoughts reported by most participants (only ~22% of all stimulus-independent thoughts), despite the long scanning session and large number of thought probes – resulting in weak statistical power. The trend-level involvement of an executive brain region, however, is intriguing and is in line with predictions we have made elsewhere (Christoff et al., 2009, Fox and Christoff, 2014, Fox et al., 2015, Ellamil et al., 2016, Fox et al., 2016). Moreover, these results are bolstered by the finding of greater grey matter concentration in numerous executive regions, including dorsal anterior cingulate cortex, in association with higher rates of intentionally directed thought (see Chapter 6).
**Neural correlates of goal-related stimulus-independent thoughts**

A wide variety of activations were observed in relation to goal-related thoughts. Especially prominent were several clusters throughout the mid-cingulate cortex extending into the premotor cortex. The mid-cingulate cortex, particularly its anterior subdivision, has a well-established role in executive control (Vogt et al., 2003) and is a key region of the frontoparietal control network (Spreng et al., 2010, Spreng et al., 2013). Such executive system recruitment is consistent with a role for these regions in planning and simulating solutions to one’s goals (Fox et al., 2015). We also observed numerous clusters in the cerebellum bilaterally. Importantly, the cerebellum is far from functionally homogenous. In addition to somatotopic mapping of motor control, a recent comprehensive meta-analysis has identified an overlapping topography related to the cerebellum’s various roles in cognition and affect (Stoodley and Schmahmann, 2009). Intriguingly, two large clusters we observed in the central portion of the cerebellum (bilaterally) overlap very closely with regions in lobules VI and VIII identified by Stoodley and Schmahmann as implicated in executive functioning (as opposed to working memory, or language, motor, somatosensory, spatial, or emotional processes). Finally, we observed clusters of activation in posterior cingulate cortex, medial visual network regions (e.g., the lingual gyrus), and a sizable cluster in left medial temporal lobe (posterior parahippocampal cortex). These regions might cooperate in creating the simulations of hypothetical scenarios presumably required in thinking about one’s goals and plans for the future (Andrews-Hanna et al., 2014).

**Neural correlates of positively-valenced thoughts**

Our major results in relation to positively-valenced thoughts were activations throughout the head, body, and tail of the left caudate nucleus. These results are consistent with many neuroimaging studies of positive affect and reward, in which basal ganglia recruitment is the
single most consistent finding (Phan et al., 2002). Intriguingly, a recent fMRI case study examined an unusual participant who claimed to experience excessive and highly-rewarding daydreaming, and found significantly increased basal ganglia and subgenual anterior cingulate recruitment in this subject during mind-wandering (Mason et al., 2013) – consistent with our results in a normal sample. Our findings in the subgenual anterior cingulate cortex area of medial prefrontal cortex are also congruent with the repeated recruitment of nearby medial prefrontal areas in rewarding emotions (Phan et al., 2002) as well as work that has specifically shown a role for nearby medial prefrontal cortex areas in affect during mind-wandering (Tusche et al., 2014).

We also observed recruitment of the right cerebellum, in the area of lobule VI. As noted above, the cerebellum is far from functionally homogenous, but rather exhibits a specific topography with respect to various cognitive and affective processes (Stoodley and Schmahmann, 2009). Importantly, the cluster of activation we observed in the right cerebellum lobule VI overlaps nearly perfectly with a meta-analytic cluster identified by Stoodley and Schmahmann associated with emotional processing (see their Fig. 1c).

As a general observation regarding positively-valenced thought, it should be noted that many of the studies included in the meta-analyses with which our results are compared here included fMRI investigations of experiences that go beyond merely ‘pleasant’ emotion. For instance, some studies also explored neural correlates related to highly-valued stimuli (i.e., ‘valuation’ studies); others involved subjects anticipating or predicting a future reward (i.e., imagining future rewards). While such studies generally involve recruitment of similar brain regions to those seen in the experience of positive affect itself, it should be noted that this neural similarity (but subjective diversity of valuation/reward experience) compounds the difficulty of interpreting our results.
We observed no significant results for negatively-valenced thoughts, but this is likely due to the very small number of such thoughts participants reported over the thought sampling session. On the other hand, the lack of any results for neutral > emotional thoughts cannot be attributed to a lack of power, as a very large number of neutral thoughts were reported. A tentative conclusion, then, is that affectively neutral thoughts simply do not lead to any brain recruitment significantly different, or greater, than emotionally-charged thoughts, but future work can hopefully examine this issue more carefully.

**Limitations**

Some limitations of the preceding study should be briefly mentioned. First, it appears that, despite our efforts to collect a very large number of thought probes, we still had too few trials to fully resolve every category of thought we examined. Specifically, the low number of negatively-valenced and intentionally-directed thoughts resulted in either marginal or no results whatsoever for the contrasts of interest. Because the experimenter naturally cannot force participants to have higher *rates* of certain kinds of self-generated thought, the only obvious solution to this problem for future work is to collect more thought probes over a longer scanning session in order to obtain more trials of each thought type and increase the statistical power of subsequent analyses. Another potential solution to this problem is to employ multivariate analysis methods; this possibility is discussed further in the subsequent section on future directions.

Another concern is ecological validity, i.e., whether thoughts reported and experienced in the MRI scanner environment are indeed representative of, and generalizable to, typical self-
generated thought in everyday life. If not, then the neural correlates of the thoughts that are experienced, too, might differ appreciably. In the subsequent Experiment 4 (Chapter 5), we go to some lengths to address this issue, and therefore refer the reader there for further discussion of this important question.

Another important limitation is that our results speak to the neural basis of specific thought types across individuals, but reveal nothing about the brain basis of interindividual differences in thought content and valence – a major area of interest in relation to clinical disorders of self-generated thought such as depressive rumination (Christoff et al., in press, DuPre and Spreng, in press). In our subsequent Experiments 5 and 6 (Chapter 6) we therefore address precisely this question.

Another major concern is the possibility of prevalent themes of thought content coloring, or even obscuring, the abstract thought qualities we were aiming to investigate here. For instance, many of the participants were graduate students, and therefore many of the thoughts reported as ‘goal-related’ are very likely to have been focused on graduate research, thesis requirements, publication of manuscripts, and so on. Especially among a relatively homogenous group such as our sample, there is the possibility that the same or very similar types of thought content could recur frequently enough to overshadow the actual abstract quality we aimed to investigate. Therefore the possibility cannot be discounted that, for instance, a significant amount of the activations observed for goal-related thoughts could in fact reflect specific thinking and planning related to writing manuscripts, theses, and so on. Two main solutions to this problem are possible: either more detailed thought content reports could be collected, and such commonalities controlled for; or a wider variety of individuals, presumably with highly disparate goals and thought content, could be recruited. The second option, however, would still not be
able to control for nearly ubiquitous human thoughts and desires, such as thought about sexual activity. If thought about sexual activity is very common across nearly all individuals, and is correlated (for instance) with positive affect, then the only way to resolve positive affect itself at the neural level, uncolored by common content themes, would be to collect more detailed thought reports and to partial out the effect due to specific thought content. Note, however, that the signal-to-noise ratio of fMRI requires a large number of each type of ‘trial’ — or thought, for our purposes. Allowing such a large number of different types of thoughts to spontaneously arise would likely require inordinate and unrealistic amounts of time spent in the MRI scanner (2-3 hours or more); resolving specific thought content and qualities at this level of detail is therefore likely a project that is at (or perhaps beyond) the boundary of what is technologically feasible with MRI.

**Conclusions and future directions**

Our results suggest that different types of self-generated thought can be successfully resolved at the neural level with fMRI using contrasts based solely on first-person reports of experience. Moreover, the activations we observed in relation to each thought type were, broadly speaking, consistent with the putative roles of these brain regions and with prior predictions about the role these areas might play in self-generated thought (Fox and Christoff, 2014, Fox et al., 2015).

The naturally low prevalence of certain thought types suggests that future work will need to collect large numbers of thought events over long scanning sessions in order to improve the discriminability of different thought types and better understand their neural basis, assuming that a general linear model (GLM) approach is used, as in the present analyses. Because of the
inherent requirement for thousands of individual contrasts using the GLM method, the
intrinsically low signal-to-noise ratio of the fMRI BOLD signal is further attenuated by the need
to control for these multiple comparisons in some way, which yet further attenuates the statistical
power of generally underpowered GLM analyses. Whereas one solution would therefore be to
collect an even larger number of trials to improve statistical power, another potential pathway to
greater statistical strength in analyzing the present dataset would be to use multivariate analysis
techniques, such as multivariate pattern analysis (Norman et al., 2006, Haxby, 2012). For
instance, one research group has used multivariate pattern analysis to decode the emotional
valence of mind-wandering episodes (Tusche et al., 2014), and others have been able with some
success to predict specific mind-wandering content from brain-wide patterns of intrinsic brain
activity as measured with fMRI (Gorgolewski et al., 2014). These methods represent powerful
techniques which might help in overcoming the lack of statistical power in certain of our
analyses (for instance for intentionally directed and negatively-valenced thoughts). We plan to
apply multivariate methods to the current dataset in future analyses intended for publication in
peer-reviewed journals.
CHAPTER 5 – STUDY 4: DOES THE MRI SCANNER ENVIRONMENT INFLUENCE THOUGHT CONTENT?

Introduction

Before discussing neural correlates of self-generated thought in any further detail, it is worth considering a (to our knowledge) still-unexplored question: does the MRI scanner environment exert a significant influence on thought content?

There are good reasons to suspect that the answer to this question is ‘yes.’ Although true adverse events are exceedingly rare (Hawkinson et al., 2012), participants entering the MRI scanner, especially for the first time, often report a wide variety of subjective effects, most notably claustrophobia and anxiety (Huettel et al., 2004). It therefore seems plausible that in a study examining the affective valence and other abstract qualities of self-generated thought, such as the present fMRI investigation (Chapter 4), thought content might be unduly influenced by reactions to the scanner environment. To our knowledge, however, this issue remains uninvestigated.

Here we sought to address this problem in a preliminary way by comparing thought content reported (i) by subjects in the MRI scanner environment with (ii) thought content reported by these same subjects but outside the scanner, seated at a desktop computer in a minimally-stimulating room; and (iii) thought content reported from this same laboratory environment in an independent sample of subjects.
Methods

Participants

Participants were the same as those employed in the fMRI experiment (Chapter 4) and the morphometric neuroimaging experiments (Chapter 6), including a total of 32 unique participants engaged in a variety of conditions (Table 18).

Table 18. Participant features for context comparisons.

<table>
<thead>
<tr>
<th>Experimental Setting</th>
<th>Sample size (N)</th>
<th>Age ± SD (years)</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI Scanner</td>
<td>16</td>
<td>28.1 ± 4.3</td>
<td>9 F, 7 M</td>
</tr>
<tr>
<td>Laboratory room (all subjects)</td>
<td>32</td>
<td>28.1 ± 4.1</td>
<td>16 F, 16 M</td>
</tr>
<tr>
<td>Laboratory room (independent of MRI sample)</td>
<td>16</td>
<td>27.8 ± 3.9</td>
<td>7 F, 9 M</td>
</tr>
<tr>
<td>Laboratory room (same sample as in MRI scanner)</td>
<td>16</td>
<td>28.1 ± 4.3</td>
<td>9 F, 7 M</td>
</tr>
</tbody>
</table>

Statistical analyses

All analyses (t-tests and correlations) were performed in SPSS Statistics 20 (Armonk, NY: IBM Corp.).
Results

Average thought content across contexts

We found that the average proportion of each type of thought across the various categories we measured was nearly identical across contexts (MRI scanner vs. laboratory environment) (Figs. 12 and 13). The only notable difference was that stimulus-independent thought tended to be somewhat higher in the MRI scanner (84.2% of thoughts) vs. in the laboratory room (75.4% of thoughts). However, this difference was much smaller when the same fMRI participants reported their thoughts in a follow-up behavioral testing session (82.5%), whereas a separate sample that only completed the laboratory testing session had a considerably lower percentage of stimulus-independent thoughts (68.7%). None of the comparisons was significant (all p’s > .05).
Figure 12 Average thought proportions inside and outside the MRI scanner. ALL Lab = thoughts reported by 32 participants in lab sessions at a desktop computer; fMRI Ss in MRI = thoughts reported in the MRI scanner by 16 participants who completed the full fMRI study; fMRI Ss in Lab = thoughts reported by the same fMRI participants in a separate behavioral testing session in the lab space at a desktop computer; Lab ONLY = thoughts averages for the participants who only completed the lab-based testing session and entered the MRI scanner only for brief anatomical scans. Note that ALL Lab = fMRI Ss in Lab + Lab ONLY.
Figure 13 Radar plot of thought proportions in the MRI scanner and in the laboratory. Details are the same as in the previous figure.

**Stability of thought content patterns at the individual level**

We then examined the stability of thought content across contexts at the individual level, i.e. whether individual participants tended to have the same proportion of, e.g., goal-related thoughts in and outside the MRI scanner (Fig. 14). Overall, we found that correlations between thought patterns at the individual level were extremely high across environmental contexts (all \(|r| (14) \geq 0.56\), all \(p\’s < 0.03\)). Note, too, that many subjects had 1-2 days between testing sessions; our data therefore demonstrate that individual thought content proportions are relatively stable over short time periods, as well. The weakest positive association was with rates of stimulus-independent (vs. stimulus-dependent) thought, although a strong correlation was still observed, \(r(14) = 0.59, p = 0.022\).
Figure 14 Stability of thought content patterns at the individual level across the MRI scanner and laboratory environments. Average, group-wide thought content changes very little across contexts (see previous two figures), and this stability is largely true at the individual level as well. Somewhat less reliable is the rate of stimulus-independent thinking ($r = .59$), but once participants are creating self-generated thought, proportions are highly stable at the individual level. The one exception to these trends is the curious reversal of proportion of positively-valenced thoughts. All relationships were statistically significant: all $|r|(14) \geq .56$, all $p$’s < .03.
Discussion

The MRI scanner environment does not appreciably influence thought content

Average proportional thought content in and outside of the MRI scanner environment were virtually identical. The only notable difference was a somewhat higher proportion of stimulus-independent thoughts in the MRI scanner as opposed to in the laboratory testing environment. This effect is probably due to the more sensory-restricted environment of the MRI scanner: in addition to general sensory impoverishment, body movement and head motion are severely restricted, preventing bottom-up attentional capture by other sensory stimuli, and although being noisy, the MRI scanner noise is repetitive and devoid of semantic content. Although we went to some lengths to ensure that our laboratory testing room was also minimally stimulating (low lighting, no objects of interest in the testing area, a curtain separating the testing area from the rest of the room, etc.), there are three sensory modalities which likely had a higher rate of meaningful input in the laboratory testing context. (i) Visual: because head movement was permitted in the laboratory room, participants were free to look around the testing area and the few objects it contained, which were still far more numerous than the single visual field of the MRI scanner (staring directly ahead at the head coil cage and reflected computer screen). (ii) Auditory: a variety of auditory stimuli from the adjacent laboratory and hallway spaces in the Brain Research Centre could not be entirely suppressed or controlled for. Although the MRI scanner also produces considerable noise, as noted above this noise is repetitive and devoid of any meaningful content, and therefore probably less likely to trigger stimulus-dependent thoughts, especially after an initial period of habituation. Conversely, overhearing conversations from outside the room seems highly likely to capture thought. (iii) Tactile: whereas participants were lying flat and instructed to move absolutely as little as possible in the MRI scanner, they
were free to move their bodies and heads in the laboratory session. Any movement would generate additional sensory feedback (e.g., the feeling of the contact of the chair the participant is sitting in, the chair moving, etc.), all of which could have triggered thoughts that would have been classified as stimulus-dependent by our criteria and instructions (since they originated from tactile inputs arising in the external environment).

Even more important, however, is that the lower rate of stimulus-independent thought in the laboratory sessions appears to be driven largely by an overall lower rate of such thought in our laboratory-only sample of subjects (purple bars). That is, our second sample of participants who only completed the laboratory session appears, by chance, to have experienced lower rates of stimulus-independent thought, and therefore lowered the ‘laboratory’ context average for all participants. These findings suggest that in fact the MRI scanner environment has a negligible effect even on rates of stimulus-independent thought, and that it was instead the heterogeneities across our relatively small groups that mainly drove this difference.

In sum, although the MRI environment is even less perceptually-stimulating than the minimalist experimental room in which the behavioral experiment took place, which helps explain the small differences in rates of stimulus-independent thought, overall the MRI scanner environment appears to have a totally negligible impact on thought content – consistent with the idea that patterns of thought are driven largely by stable individual trait-like tendencies. Aside from lending support to the validity of our own fMRI results (Chapter 4), these findings add to the ecological validity of many previous fMRI investigations of self-generated thought (reviewed in Fox et al., 2015).
Individual patterns of thought content are stable over time and across contexts

A second finding of our analysis is that individual patterns of thought are highly stable across contexts (MRI scanner vs. laboratory) and short time periods (1-3 days). The first exception to this was the lower correlation between rates of stimulus-independent thought itself, which seems to vary more widely, although the correlation was still considerable ($r = .56$). The other main exception was positively-valenced thought which, strangely, showed a significant negative correlation ($r = -.56$) – in stark contrast to every other thought category, which showed much stronger and positive correlations (all $r$’s $\geq .84$) across contexts. Despite the exception of positively-valenced thought, overall our results suggest that a ~45 min period consisting of 120 thought probes captures relatively stable individual tendencies in thought content that are largely invariable across contexts and over short time periods. If thought content reported in such a session is indeed representative of trait-like individual differences, then these reports might have validity as individual difference measures that could be associated with individual neuroanatomical heterogeneities in grey and white matter structure – a hypothesis we examine in the following chapter and Experiments.
CHAPTER 6 – STUDIES 5 AND 6: INVESTIGATING THE RELATIONSHIP BETWEEN INDIVIDUAL DIFFERENCES IN SELF-GENERATED THOUGHT AND BRAIN STRUCTURE USING MORPHOMETRIC NEUROIMAGING

Introduction: the relevance of interindividual neuroanatomical heterogeneity

While plasticity in the adult mammalian brain was once thought to be absolutely minimal, if not impossible (Ramón y Cajal, 1928), structural plasticity at the microstructural scale (dendritic spine density, axon process length, etc.) is now well documented in the fully mature brains of non-mammalian species, rodents, and non-human primates. Although investigation of adult structural brain heterogeneities remains restricted (due to the limitations of current neuroimaging methodologies) to the mesostructural level in humans (e.g., grey matter volume and concentration, fractional anisotropy, etc.), nonetheless wide evidence is now available for functionally-relevant neuroanatomical heterogeneity in the adult brain (May and Gaser, 2006, Zatorre et al., 2012). Perhaps more impressive still has been the discovery of neurogenesis (growth, differentiation, and proliferation of new neurons) in the adult brain, now observed in song birds (Goldman and Nottebohm, 1983), rodents (Altman and Das, 1965), and a variety of other mammalian species (Fuchs and Gould, 2000) – even humans (Eriksson et al., 1998).

Such changes in brain structure are typically associated with overt physical practice, training, environmental enrichment, and the like, but significant associations have also been found between brain morphology and individual trait variables in both pathological and normal populations. For instance, systematic differences in brain morphology have been reported in large-scale meta-analyses of schizophrenia (Antonova et al., 2004), major depression
(Koolschijn et al., 2009), and obsessive-compulsive disorder (Radua and Mataix-Cols, 2009). Morphometric neuroimaging has also revealed neuroanatomical differences associated with less marked differences in personality and behavior, however, such as scores on the ‘Big Five’ personality traits (DeYoung et al., 2010) as well as other measures of personality or cognitive style (Kaasinen et al., 2005, Gardini et al., 2009).

Researchers, particularly the pioneering daydreaming researcher Jerome Singer, have long suggested that differences in self-generated thought content might reflect not only the momentary contents of the mind, but differing *trait* styles of thinking (Singer and Brown, 1960, Singer, 1961, Singer and McCraven, 1961, Singer and McCraven, 1962). Based on results from numerous studies using their *Imaginal Process Inventory* questionnaire, which assessed many different dimensions of self-generated thought, Singer, Antrobus, Giambra, and their colleagues argued for the existence of three broad styles of thinking: *positive-constructive*, *guilty-dysphoric*, and *guilty-fear-of-failure* (Singer and Antrobus, 1970, 1972, Giambra, 1974, 1978, 1980). Follow-up work also found intuitively appealing associations between these thinking styles and scores on Big Five personality traits: for instance, *positive-constructive* daydreaming was significantly associated with Openness to Experience, whereas *guilty-dysphoric* daydreaming was associated with Neuroticism (Zhiyan and Singer, 1997).

It therefore seemed plausible to us that, as with other normal and clinical differences in personality, thinking, and behavior, major differences in patterns of self-generated thought might be associated with corresponding differences in grey and white matter structure in the central nervous system. Before describing our own study, we first review the results of a single pioneering study in this domain.
**Previous research using morphometric neuroimaging to investigate self-generated thought**

To our knowledge, only a single study to date has investigated aspects of self-generated thought with morphometric neuroimaging (Bernhardt et al., 2014). Bernhardt and colleagues (2014) investigated whether the frequency of task-unrelated thought during either a simple or demanding cognitive task was associated with differences in cortical thickness throughout the brain. During execution of the tasks, participants were periodically given experience sampling probes asking whether their current thoughts were related or unrelated to the task at hand. The authors found that the frequency of task-unrelated thought during the simple, undemanding task predicted increased cortical thickness in a large swathe of medial prefrontal cortex, extending into the anterior and mid-cingulate cortex. An additional large cluster was observed on the left lateral surface, extending from middle frontal regions posteriorly toward the temporoparietal junction. Sub-threshold (trend) clusters were also observed in rostrolateral and dorsolateral prefrontal cortex. These results suggested to the authors a potential neuroanatomical basis of individuals’ tendency to decouple from the here and now and instead engage in self-generated forms of thought, and are consistent with a growing body of fMRI research showing the importance of medial prefrontal regions to self-generated thought (Fox et al., 2015).

Several major limitations of this study encourage caution when interpreting its results, however. The most important caveat is the very small number of thought probes employed (an average of only ~7 probes in each task, for each participant). That is, on the basis of only 7 thought probes administered over a ~7 min task, the authors constructed individual difference scores representing the proportion of task-unrelated thought, which they then used to predict cortical thickness throughout the brain. Clearly, the potential for temporary state effects to influence responses to just 7 thought probes over ~7 min is a serious limitation. The same
concern can of course be raised with respect to our own methods, although we employed 120 thought probes over ~45 min, somewhat mitigating these concerns (see the Discussion for detailed arguments on this point). The other major limitation is that only a single dimension of self-generated thought was assessed: its frequency. Although interesting, this study left unexplored the many potentially intriguing individual differences in the particular content and valence of self-generated thought, which we sought to explore in our investigation.

The need for further morphometric neuroimaging of self-generated thought

The accumulated evidence that styles of thinking may be individually distinctive and stable, together with the dearth of research examining neuroanatomical correlates of thought content, suggests the need for more research along these lines. Moreover, as noted above, the one study that has conducted such an investigation (Bernhardt et al., 2014), used very few thought probes and investigated only the frequency of task-unrelated thoughts (but not their content).

We therefore designed a study to explore neuroanatomical associations with various other dimensions of self-generated thought using a much larger number of thought probes (n = 120). We computed individual difference scores for each of our participants based on their responses to 120 thought probes over a ~1 hr period, and used these scores as predictors of grey matter concentration and fractional anisotropy (fractional anisotropy). Because our thought probes assessed multiple dimensions of self-generated thought, we were able to explore several dimensions simultaneously, including the rate of stimulus-independent thought, the sense of agency or control over thoughts, the tendency of thoughts to be related to one’s goals or current concerns, and thoughts’ affective valence.
Methods

Participants

In addition to the 16 participants who engaged in the fMRI study (Chapter 4), 18 more participants were recruited for the morphometric neuroimaging study, in order to increase sample size to a level capable of yielding reasonably stable correlation coefficients (i.e., \( N > \sim 25 \); (Glass and Hopkins, 1970)). Data collection is ongoing to allow for a much better-powered sample of approximately \( N = 50 \) (Yarkoni, 2009). One participant’s entire dataset was corrupted due to an error at the MRI scanner, and another was excluded from analyses due to a severe paucity of stimulus-independent thoughts (\( n = 12 \), out of 120 thought probes), which precluded any stable individual measures of particular thought types for this participant. In total, then, 32 participants (16 female; mean age in years ± SD = 28 ± 4.1) completed the behavioral thought-sampling protocol outside of the scanner and had artifact-free 3D T1-weighted anatomical images suitable for voxel-based morphometric analyses of grey matter concentration (Experiment 5). Two further participants had susceptibility artifacts in their DTI scans and were removed from white matter analyses, leaving a total sample for the DTI investigation (Experiment 6) of 30 participants (14 female; mean age in years ± SD = 28 ± 4.0).

Most participants were graduate students at the University of British Columbia, recruited via snowball (i.e., word-of-mouth referral) sampling. All participants selected for participation in the study were neurologically and psychiatrically healthy; right-handed (to minimize any functional or structural neuroanatomical differences due to language lateralization); and provided written informed consent in accordance with the ethical standards set by the Clinical Research Ethics Board (CREB) of the University of British Columbia, Vancouver, Canada. The study
protocol was reviewed and approved by the UBC MRI Research Centre. Participants were compensated $20/hr for their time.

*Procedure and behavioral thought-sampling session*

All participants spent approximately 1 hr completing the behavioral portion of the experiment in laboratory space at the Brain Research Centre (BRC) inside the UBC Hospital. E-Prime 2 was used to present the thought probes (Psychology Software Tools Inc., Sharpsburg, PA). Upon arrival at the lab space, participants completed a consent form, and were then given a detailed explanation of the study and the thought probe questions by the experimenter. A scripted explanation was used to provide identical examples of each thought type for every participant. Participants were then left alone to engage in two practice blocks (24 probes each) of the study (~15 min), during which they had an unlimited time to classify their thoughts and respond to each probe. Following the practice blocks, the experimenter discussed the experiment and thought probe questions further with each participant in order to clarify any ambiguities in the questions or in the way in which the individual’s thoughts should be classified. Because each probe had a limited number of discrete responses (indeed, response options were mostly binarized), it was emphasized to participants that they should in every case choose the *best* answer, acknowledging that in some cases thoughts might fall somewhere in between the two response options provided and might not be easily classifiable.

When all questions had been clarified, the participant was left alone in a minimally-stimulating room with a curtain drawn to complete the full thought-sampling task, seated at a desktop computer. The thought sampling experiment was divided into 5 blocks of 24 probes each, for a total of 120 probes. Each block proceeded automatically, with 8 s allowed for each
response to each probe question, but participants had the option of taking brief, self-paced breaks between each block. Upon completion of the experiment, participants recalled the experimenter (who waited outside in an adjacent room), were debriefed, and were compensated $20/hr for their time. Most participants then proceeded directly to the MRI scanner for anatomical scanning (3D T1 and DTI scans); due to some scheduling conflicts, some participants were instead scanned within 1-2 days of their behavioral testing session.

*Calculation of individual difference scores in self-generated thought*

Though responses from each participant were collated into a database and then converted to percentages (proportions) to control for differences in the absolute number of thoughts reported by each participant. This resulted in individual scores for the percentage of stimulus-independent vs. stimulus-dependent thoughts; spontaneously unfolding vs. intentionally directed thoughts; goal-related vs. goal-unrelated thoughts; and positively-valenced, negatively-valenced, or neutral thoughts. All individual scores are reported in the Results section, below.

*MRI data acquisition*

Imaging was conducted at the UBC MRI Research Centre (Vancouver, British Columbia). MRI data were collected using a 3.0-Tesla Philips Intera MRI scanner (Best, Netherlands) with a standard 8-element 6-channel phased array head coil with parallel imaging capability (SENSE; (Pruessmann et al., 1999)). Head movement was restricted using foam padding around the head. MRI scanning was conducted either on the same day that the participants completed the behavioral training and thought-sampling session, or within 1-2 days if same-day scheduling was not possible.
For each subject, we collected (a) a high-resolution 3D T1-weighted anatomical volume (SPGR: TR = 2 s; TE = 3.53 ms; 175 interleaved axial slices covering the whole brain, 1 mm thick with 0 mm skip; FOV: 256 × 256 × 175 mm; matrix size: 256 × 250; 1 x 1 x 1 mm³ isotropic voxels). We also collected (b) a diffusion-weighted anatomical volume using high angular resolution diffusion-weighted imaging (HARDI) (Tuch et al., 2002). The DTI acquisition protocol included a whole-brain sequence with 1 non-diffusion-weighted image and 60 directionally sensitized diffusion-weighted images, with the 61 brain slices oriented obliquely to the anterior commissure–posterior commissure line, 2.2 mm thick with 0 mm skip (TR = 7909 ms, TE = 60 ms, flip angle: 90°, FOV: 224 × 154 × 224 mm, matrix size: 100 × 100, voxel dimensions: 2.5 mm × 2.5 mm × 2.5 mm). All image data were visually inspected for apparent artifacts due to subject motion and instrumental malfunction; two subjects were removed from white matter analyses due to susceptibility artifacts in their DTI data.

**Preprocessing of 3D T1 anatomical images**

All preprocessing was performed offline using SPM8 statistical parametric mapping software package (Wellcome Department of Imaging Neuroscience, London). Each subject’s high-resolution structural image was segmented into multiple tissue types, including grey matter, white matter, and cerebrospinal fluid. A subject-specific brain template was then created using the DARTEL tool in SPM8 (Ashburner, 2007) to determine the nonlinear deformations required for warping all grey and white matter images to match one another. The subject-specific mean template was then nonlinearly normalized to MNI space. Lastly, an 8-mm full-width at half maximum (FWHM) Gaussian kernel was applied to spatially smooth the data.
**Preprocessing of DTI images**

Eddy current and head motion correction were done via a linear, affine registration (FLIRT; Jenkinson and Smith, 2001, Jenkinson et al., 2002, Greve and Fischl, 2009). Non-brain voxels were removed using FSL’s brain extraction tool (Smith, 2002). DTI eigenvalues (where $\lambda_1$ is along the principal diffusion axis and $\lambda_2$ and $\lambda_3$ are orthogonal to $\lambda_1$) were calculated using software in FSL’s diffusion toolbox (Smith et al., 2004, Jenkinson et al., 2012), from which the relevant DTI parameter (fractional anisotropy) was derived. To identify regions of interest (ROIs), the FA maps of all subjects were registered into the Johns Hopkins University (JHU) 1 mm fractional anisotropy (FA) standard brain template space (Wakana et al., 2007, Hua et al., 2008) via FLIRT (12 degrees of freedom, cost function = correlation ratio) and the nonlinear registration tool FNIRT (Andersson et al., 2007b, a), which uses a b-spline representation of the registration warp field (Rueckert et al., 1999) (Jacobian range = 0.1–10).

**Voxel-based morphometry analysis of grey matter concentration**

Voxel-based morphometry (VBM) is a class of techniques for automatically assessing neuroanatomical heterogeneities in grey matter, either between groups or longitudinally in the same subjects (Ashburner and Friston, 2000, Mechelli et al., 2005). Following the preprocessing steps described above, our various individual difference scores were entered into a multiple regression design matrix in SPM8 and correlated with grey matter intensity at each voxel in whole-brain analyses across our sample ($N = 32$). Because both grey and white matter are known to be negatively impacted by normal aging (Good et al., 2002), age was entered as a covariate of no interest in our analyses (despite the relatively narrow range and low variance of ages in our
sample) to control for any effects due to normal age-related atrophy. Sex was also entered as a covariate of no interest.

*Tract-based analysis of fractional anisotropy*

In order to assess fractional anisotropy, we first defined 20 white matter tract regions-of-interest (ROIs) based on the Johns Hopkins University (JHU) probabilistic white matter tractography atlas (Fig. 15) (Wakana et al., 2007, Hua et al., 2008). This procedure has numerous benefits as opposed to a whole-brain white matter analysis: it excludes all voxel tissue types (e.g., grey matter and cerebrospinal fluid) of no interest to the analyses by creating a white matter tract-exclusive mask, therefore precluding spurious results; it reduces the need for thousands of multiple comparisons at every white matter voxel in the brain, thereby greatly increasing statistical power; and it enhances the interpretability of results by confining findings to major tracts as a whole, since specific functional roles are mostly unknown for particular loci along white matter tracts. Mean FA was calculated for each of the 20 white matter tracts for each participant, and these values were correlated with individual differences scores on thought content.

Because of our small sample size (N = 30), and the additional loss of degrees of freedom when controlling for age and sex, we considered this Experiment merely a pilot study and therefore used liberal statistical thresholds. Given the recent criticisms of null hypothesis significance testing and arbitrary p-value thresholds (Cumming, 2012, 2013), we sought to report potentially practically meaningful effect sizes (which are unaffected by sample size) rather than simply searching for statistically ‘significant’ findings (which are highly dependent on sample size). As a criterion for reporting of preliminary results, we therefore set a ‘medium’ effect size
(r ≅ .25-.35) as defined by Cohen (Cohen, 1992), i.e., a minimum of r > .25 for our partial correlation coefficients.

**Figure 15.** White matter tract regions-of-interest (ROIs). Mean fractional anisotropy (a standard measure of fractional anisotropy) was calculated for each tract, based on the Johns Hopkins University probabilistic white matter tract atlas (Wakana et al., 2007, Hua et al., 2008), and correlated with individual differences in thought content. ROI image reproduced from Luders et al. (2011).
Behavioral results: individual differences in thought content

All thought probe responses were condensed into individual percentage scores for each thought type, as described in the Methods. Results are reported in Table 19.

Table 19. Individual patterns of thought content

<table>
<thead>
<tr>
<th>Subject</th>
<th>% SIT</th>
<th>% Directed</th>
<th>% Related</th>
<th>% (–)</th>
<th>% Neutral</th>
<th>% (+)</th>
<th>Age</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>S03</td>
<td>81.74</td>
<td>17.02</td>
<td>68.09</td>
<td>30.85</td>
<td>23.40</td>
<td>45.74</td>
<td>22</td>
<td>M</td>
</tr>
<tr>
<td>S06</td>
<td>79.82</td>
<td>30.23</td>
<td>30.59</td>
<td>12.79</td>
<td>61.63</td>
<td>25.58</td>
<td>24</td>
<td>F</td>
</tr>
<tr>
<td>S07*</td>
<td>95.00</td>
<td>18.42</td>
<td>24.11</td>
<td>30.09</td>
<td>33.63</td>
<td>36.28</td>
<td>34</td>
<td>F</td>
</tr>
<tr>
<td>S08</td>
<td>88.14</td>
<td>1.94</td>
<td>36.27</td>
<td>26.21</td>
<td>41.75</td>
<td>32.04</td>
<td>35</td>
<td>F</td>
</tr>
<tr>
<td>S09</td>
<td>82.35</td>
<td>2.38</td>
<td>8.54</td>
<td>8.43</td>
<td>72.29</td>
<td>19.28</td>
<td>28</td>
<td>M</td>
</tr>
<tr>
<td>S10</td>
<td>71.17</td>
<td>33.77</td>
<td>84.62</td>
<td>5.33</td>
<td>40.00</td>
<td>54.67</td>
<td>33</td>
<td>M</td>
</tr>
<tr>
<td>S11</td>
<td>76.99</td>
<td>41.38</td>
<td>45.98</td>
<td>14.94</td>
<td>43.68</td>
<td>41.38</td>
<td>28</td>
<td>F</td>
</tr>
<tr>
<td>S12</td>
<td>76.79</td>
<td>43.02</td>
<td>58.82</td>
<td>33.73</td>
<td>31.33</td>
<td>34.94</td>
<td>35</td>
<td>F</td>
</tr>
<tr>
<td>S13</td>
<td>89.29</td>
<td>20.00</td>
<td>22.67</td>
<td>16.00</td>
<td>66.67</td>
<td>17.33</td>
<td>24</td>
<td>F</td>
</tr>
<tr>
<td>S14</td>
<td>89.19</td>
<td>62.63</td>
<td>54.55</td>
<td>6.06</td>
<td>72.73</td>
<td>21.21</td>
<td>26</td>
<td>F</td>
</tr>
<tr>
<td>S15</td>
<td>85.71</td>
<td>1.96</td>
<td>50.98</td>
<td>10.78</td>
<td>67.65</td>
<td>21.57</td>
<td>26</td>
<td>F</td>
</tr>
<tr>
<td>S16</td>
<td>86.67</td>
<td>75.73</td>
<td>79.81</td>
<td>9.62</td>
<td>39.42</td>
<td>50.96</td>
<td>24</td>
<td>F</td>
</tr>
<tr>
<td>S17</td>
<td>94.92</td>
<td>1.79</td>
<td>88.18</td>
<td>38.39</td>
<td>6.25</td>
<td>55.36</td>
<td>26</td>
<td>F</td>
</tr>
<tr>
<td>S18</td>
<td>86.24</td>
<td>40.43</td>
<td>59.14</td>
<td>16.30</td>
<td>32.61</td>
<td>51.09</td>
<td>28</td>
<td>M</td>
</tr>
<tr>
<td>S19</td>
<td>67.02</td>
<td>29.51</td>
<td>28.57</td>
<td>9.68</td>
<td>87.10</td>
<td>3.23</td>
<td>31</td>
<td>M</td>
</tr>
<tr>
<td>S20</td>
<td>57.39</td>
<td>4.62</td>
<td>11.86</td>
<td>30.00</td>
<td>35.00</td>
<td>35.00</td>
<td>26</td>
<td>M</td>
</tr>
<tr>
<td>S21</td>
<td>57.76</td>
<td>9.09</td>
<td>20.90</td>
<td>24.24</td>
<td>48.48</td>
<td>27.27</td>
<td>25</td>
<td>F</td>
</tr>
<tr>
<td>S22*</td>
<td>39.13</td>
<td>6.67</td>
<td>38.64</td>
<td>22.22</td>
<td>28.89</td>
<td>48.89</td>
<td>31</td>
<td>F</td>
</tr>
<tr>
<td>S24</td>
<td>82.29</td>
<td>27.27</td>
<td>52.11</td>
<td>34.25</td>
<td>26.03</td>
<td>39.73</td>
<td>26</td>
<td>F</td>
</tr>
<tr>
<td>S25</td>
<td>82.11</td>
<td>28.95</td>
<td>56.52</td>
<td>22.97</td>
<td>52.70</td>
<td>24.32</td>
<td>27</td>
<td>M</td>
</tr>
<tr>
<td>S26</td>
<td>57.50</td>
<td>37.31</td>
<td>32.35</td>
<td>20.29</td>
<td>43.48</td>
<td>36.23</td>
<td>37</td>
<td>M</td>
</tr>
<tr>
<td>S27</td>
<td>73.87</td>
<td>21.79</td>
<td>20.78</td>
<td>2.47</td>
<td>34.57</td>
<td>62.96</td>
<td>27</td>
<td>M</td>
</tr>
<tr>
<td>S28</td>
<td>65.83</td>
<td>24.68</td>
<td>34.62</td>
<td>30.38</td>
<td>60.76</td>
<td>8.86</td>
<td>27</td>
<td>M</td>
</tr>
<tr>
<td>S29</td>
<td>63.46</td>
<td>33.33</td>
<td>90.91</td>
<td>15.38</td>
<td>32.31</td>
<td>52.31</td>
<td>30</td>
<td>M</td>
</tr>
<tr>
<td>S30</td>
<td>86.84</td>
<td>44.90</td>
<td>47.47</td>
<td>30.93</td>
<td>43.30</td>
<td>25.77</td>
<td>21</td>
<td>M</td>
</tr>
<tr>
<td>S31</td>
<td>97.48</td>
<td>50.45</td>
<td>58.62</td>
<td>15.52</td>
<td>39.66</td>
<td>44.83</td>
<td>23</td>
<td>M</td>
</tr>
<tr>
<td>S32</td>
<td>50.00</td>
<td>14.04</td>
<td>36.84</td>
<td>19.64</td>
<td>50.00</td>
<td>30.36</td>
<td>26</td>
<td>F</td>
</tr>
<tr>
<td>S33</td>
<td>60.00</td>
<td>37.88</td>
<td>40.00</td>
<td>39.39</td>
<td>21.21</td>
<td>39.39</td>
<td>30</td>
<td>F</td>
</tr>
<tr>
<td>S34</td>
<td>72.97</td>
<td>39.51</td>
<td>55.00</td>
<td>28.75</td>
<td>17.50</td>
<td>53.75</td>
<td>28</td>
<td>M</td>
</tr>
<tr>
<td>S35</td>
<td>69.75</td>
<td>13.58</td>
<td>46.34</td>
<td>26.25</td>
<td>32.50</td>
<td>41.25</td>
<td>26</td>
<td>M</td>
</tr>
<tr>
<td>S36</td>
<td>23.15</td>
<td>20.83</td>
<td>25.00</td>
<td>32.00</td>
<td>28.00</td>
<td>40.00</td>
<td>31</td>
<td>M</td>
</tr>
<tr>
<td>S37</td>
<td>93.75</td>
<td>2.97</td>
<td>93.33</td>
<td>30.48</td>
<td>32.38</td>
<td>37.14</td>
<td>34</td>
<td>F</td>
</tr>
</tbody>
</table>

SIT: stimulus-independent thought. Asterisks (*) denote subjects not included in DTI analyses due to susceptibility artifacts in diffusion-weighted data.
Morphometric neuroimaging results I: grey matter concentration

Grey matter concentration associated with frequency of intentionally directed stimulus-independent thought

We found a large number of clusters of grey matter concentration throughout the brain associated with individual tendency toward intentionally directed thoughts. Most striking were numerous clusters in prefrontal executive regions, including dorsolateral prefrontal cortex bilaterally (BA 46/9/10), bilateral subgenual anterior cingulate cortex/orbitofrontal cortex (BA 25/12), and right ventrolateral prefrontal cortex (BA 47). Also salient were clusters in bilateral posterior parahippocampal/hippocampal cortex (BA 30) and left superior temporal gyrus/anterior insula (BA 22/13). Results are summarized in Table 20 and Fig. 16.
Table 20. Grey matter concentration peaks for higher frequency of intentionally directed thoughts

<table>
<thead>
<tr>
<th>Region</th>
<th>Side</th>
<th>BA</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxels</th>
<th>t-statistic</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superior temporal gyrus/insula</td>
<td>L</td>
<td>22/13</td>
<td>-50</td>
<td>1</td>
<td>-4</td>
<td>535</td>
<td>5.78</td>
<td>4.65</td>
</tr>
<tr>
<td>Subgenual cingulate cortex/orbitofrontal cortex</td>
<td>L</td>
<td>25/12</td>
<td>-20</td>
<td>10</td>
<td>-16</td>
<td>575</td>
<td>5.04</td>
<td>4.22</td>
</tr>
<tr>
<td>Dorsolateral prefrontal cortex</td>
<td>L</td>
<td>46/9</td>
<td>-54</td>
<td>31</td>
<td>20</td>
<td>131</td>
<td>4.86</td>
<td>4.10</td>
</tr>
<tr>
<td>Lingual gyrus</td>
<td>R</td>
<td>19</td>
<td>21</td>
<td>-59</td>
<td>-1</td>
<td>355</td>
<td>4.62</td>
<td>3.95</td>
</tr>
<tr>
<td>Posterior parahippocampal cortex/hippocampus</td>
<td>L</td>
<td>30</td>
<td>-30</td>
<td>-33</td>
<td>-4</td>
<td>751</td>
<td>4.54</td>
<td>3.90</td>
</tr>
<tr>
<td>Rostrolateral/dorsolateral prefrontal cortex</td>
<td>R</td>
<td>10/46</td>
<td>39</td>
<td>39</td>
<td>-1</td>
<td>93</td>
<td>4.30</td>
<td>3.74</td>
</tr>
<tr>
<td>Ventrolateral prefrontal cortex</td>
<td>R</td>
<td>47</td>
<td>43</td>
<td>22</td>
<td>-9</td>
<td>72</td>
<td>4.22</td>
<td>3.68</td>
</tr>
<tr>
<td>Subgenual cingulate cortex/orbitofrontal cortex</td>
<td>L</td>
<td>25</td>
<td>-2</td>
<td>3</td>
<td>-13</td>
<td>117</td>
<td>4.19</td>
<td>3.66</td>
</tr>
<tr>
<td>Posterior parahippocampal cortex/hippocampus</td>
<td>R</td>
<td>30</td>
<td>22</td>
<td>-39</td>
<td>5</td>
<td>81</td>
<td>3.89</td>
<td>3.45</td>
</tr>
<tr>
<td>Middle occipital gyrus</td>
<td>R</td>
<td>19</td>
<td>39</td>
<td>-78</td>
<td>20</td>
<td>7</td>
<td>3.87</td>
<td>3.43</td>
</tr>
<tr>
<td>Superior temporal gyrus</td>
<td>R</td>
<td>21</td>
<td>48</td>
<td>0</td>
<td>-15</td>
<td>44</td>
<td>3.82</td>
<td>3.40</td>
</tr>
<tr>
<td>Putamen</td>
<td>L</td>
<td>–</td>
<td>-30</td>
<td>2</td>
<td>8</td>
<td>21</td>
<td>3.80</td>
<td>3.38</td>
</tr>
<tr>
<td>Precuneus</td>
<td>R</td>
<td>7</td>
<td>20</td>
<td>-44</td>
<td>56</td>
<td>16</td>
<td>3.72</td>
<td>3.33</td>
</tr>
<tr>
<td>Premotor cortex</td>
<td>R</td>
<td>6</td>
<td>18</td>
<td>0</td>
<td>-15</td>
<td>20</td>
<td>3.71</td>
<td>3.32</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>L</td>
<td>–</td>
<td>-36</td>
<td>33</td>
<td>51</td>
<td>12</td>
<td>3.67</td>
<td>3.29</td>
</tr>
<tr>
<td>Dorsal anterior cingulate cortex/premotor cortex</td>
<td>L</td>
<td>32/6</td>
<td>-10</td>
<td>21</td>
<td>42</td>
<td>35</td>
<td>3.65</td>
<td>3.27</td>
</tr>
<tr>
<td>Lateral globus pallidus</td>
<td>L</td>
<td>–</td>
<td>-16</td>
<td>0</td>
<td>-4</td>
<td>6</td>
<td>3.52</td>
<td>3.17</td>
</tr>
</tbody>
</table>

Height threshold $t = 3.41$ ($p < .001$, uncorrected), extent threshold $k > 5$. 
Figure 16. Grey matter concentration peaks for higher frequency of intentionally directed thoughts. Height threshold $t = 2.76$ ($p < .005$, uncorrected), extent threshold $k > 5$. aMCC: anterior mid-cingulate cortex; BA: Brodmann area; dACC: dorsal anterior cingulate cortex; DLPFC: dorsolateral prefrontal cortex; Ins: insula; mOcc: medial occipital cortex; MTL: medial temporal lobe; PCC: posterior cingulate cortex; PMC: premotor cortex; Prec: precuneus; sgCC: subgenual cingulate cortex; STG: superior temporal gyrus.
Grey matter concentration associated with frequency of goal-related stimulus-independent thought

For correlations between grey matter concentration and frequency of goal-related thoughts, we found two significant clusters, in the posterior cingulate cortex extending anteriorly into the left thalamus, and in the left parahippocampal cortex (BA 35), extending superiorly into the thalamus. At a slightly relaxed threshold used for visualization purposes ($p < .005$, uncorrected), we observed other clusters of correlation as well in dorsal anterior cingulate cortex anterior mid-cingulate cortex (BA 24), left dorsolateral/rostralateral prefrontal cortex (BA 10/9), left inferior parietal lobule (BA 40), and in right parahippocampal cortex (BA 35/36). Results are summarized in Table 21 and Fig. 17.

Table 21. Grey matter concentration peaks for higher frequency of goal-related thoughts

<table>
<thead>
<tr>
<th>Region</th>
<th>Side</th>
<th>BA</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxels</th>
<th>$t$-statistic</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Posterior cingulate cortex/thalamus</td>
<td>M</td>
<td>30/29/26</td>
<td>-4</td>
<td>-33</td>
<td>12</td>
<td>266</td>
<td>4.17</td>
<td>3.65</td>
</tr>
<tr>
<td>Parahippocampal cortex/thalamus</td>
<td>L</td>
<td>35</td>
<td>-16</td>
<td>-32</td>
<td>-4</td>
<td>30</td>
<td>4.00</td>
<td>3.53</td>
</tr>
</tbody>
</table>

Height threshold $t = 3.41$ ($p < .001$, uncorrected), extent threshold $k > 5$. 
Figure 17. Grey matter concentration peaks for greater goal-related thoughts. Height threshold \( t = 2.76 \) \((p < .005, \text{ uncorrected})\), extent threshold \( k > 5 \). aMCC: anterior mid-cingulate cortex; dACC: dorsal anterior cingulate cortex; DLPFC: dorsolateral prefrontal cortex; IPL: inferior parietal lobule; L: left; PCC: posterior cingulate cortex; PHC: parahippocampal cortex; R: right; Thal: thalamus.
Grey matter concentration associated with frequency of positively-valenced stimulus-independent thought

We then examined correlations between grey matter concentration and frequency of positively-valenced thoughts. We found significant associations in a large swathe of anterior mid-cingulate cortex (BA 24), in posterior mid-cingulate cortex (BA 24), bilaterally in the inferior parietal lobule (BA 40), and in the right inferior temporal gyrus (BA 20). See Table 22 and Fig. 18.

Table 22. Grey matter concentration peaks for higher frequency of positively-valenced thoughts

<table>
<thead>
<tr>
<th>Region</th>
<th>Side</th>
<th>BA</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxels</th>
<th>( t )-statistic</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anterior mid-cingulate cortex</td>
<td>M</td>
<td>24</td>
<td>2</td>
<td>10</td>
<td>27</td>
<td>96</td>
<td>4.11</td>
<td>3.60</td>
</tr>
<tr>
<td>Inferior parietal lobule</td>
<td>L</td>
<td>40</td>
<td>-51</td>
<td>-33</td>
<td>36</td>
<td>21</td>
<td>4.05</td>
<td>3.56</td>
</tr>
<tr>
<td>Posterior mid-cingulate cortex</td>
<td>M</td>
<td>24</td>
<td>4</td>
<td>-6</td>
<td>30</td>
<td>34</td>
<td>3.86</td>
<td>3.43</td>
</tr>
<tr>
<td>Inferior parietal lobule</td>
<td>L</td>
<td>40</td>
<td>-46</td>
<td>-33</td>
<td>51</td>
<td>50</td>
<td>3.82</td>
<td>3.40</td>
</tr>
<tr>
<td>Inferior temporal gyrus</td>
<td>R</td>
<td>20</td>
<td>54</td>
<td>-26</td>
<td>-19</td>
<td>6</td>
<td>3.54</td>
<td>3.19</td>
</tr>
</tbody>
</table>

Height threshold \( t = 3.41 \) (\( p < .001 \), uncorrected), extent threshold \( k > 5 \).
Figure 18. Grey matter concentration peaks for greater positively-valenced thoughts. Height threshold \( t = 2.76 \) \((p < .005\), uncorrected\), extent threshold \( k > 5 \). IPL: inferior parietal lobule; ITG: inferior temporal gyrus; LG: lingual gyrus; MCC: mid-cingulate cortex; PHC: parahippocampal cortex.
Grey matter concentration associated with frequency of negatively-valenced stimulus-independent thought

We found several clusters, mostly spread throughout premotor cortices, as well as in the left cerebellum and precuneus. Results are summarized in Table 23 and Fig. 19.

Table 23. Grey matter concentration peaks for higher frequency of negatively-valenced thoughts

<table>
<thead>
<tr>
<th>Region</th>
<th>Side</th>
<th>BA</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxels</th>
<th>t-statistic</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Premotor cortex</td>
<td>R</td>
<td>6/32</td>
<td>20</td>
<td>18</td>
<td>44</td>
<td>215</td>
<td>4.95</td>
<td>4.16</td>
</tr>
<tr>
<td>Precuneus</td>
<td>L</td>
<td>7</td>
<td>-12</td>
<td>-66</td>
<td>51</td>
<td>32</td>
<td>3.94</td>
<td>3.48</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>M</td>
<td>-</td>
<td>-9</td>
<td>-71</td>
<td>-25</td>
<td>57</td>
<td>3.78</td>
<td>3.37</td>
</tr>
<tr>
<td>Premotor cortex</td>
<td>R</td>
<td>6/32</td>
<td>24</td>
<td>18</td>
<td>71</td>
<td>24</td>
<td>3.77</td>
<td>3.36</td>
</tr>
<tr>
<td>Premotor cortex</td>
<td>R</td>
<td>6</td>
<td>18</td>
<td>7</td>
<td>59</td>
<td>8</td>
<td>3.73</td>
<td>3.33</td>
</tr>
<tr>
<td>Premotor cortex</td>
<td>M</td>
<td>6</td>
<td>2</td>
<td>24</td>
<td>57</td>
<td>9</td>
<td>3.55</td>
<td>3.20</td>
</tr>
<tr>
<td>Premotor cortex</td>
<td>M</td>
<td>6</td>
<td>10</td>
<td>13</td>
<td>62</td>
<td>7</td>
<td>3.53</td>
<td>3.18</td>
</tr>
</tbody>
</table>

Height threshold \( t = 3.41 \) (\( p < .001 \), uncorrected), extent threshold \( k > 5 \).

Figure 19. Grey matter concentration associated with frequency of negatively-valenced thought. Height threshold \( t = 2.76 \) (\( p < .005 \), uncorrected), extent threshold \( k > 5 \). PMC: premotor cortex; Prec: precuneus.
Morphometric neuroimaging results II: fractional anisotropy

As discussed in the Methods, above, we conducted a preliminary pilot study of the relationship between mean fractional anisotropy (FA) and various individual thought tendencies in white matter tracts throughout the brain. As this was an exploratory analysis, we merely report relationships that showed medium-sized ($r = .25$) effects or greater. As tendency toward both positive and negative emotion showed essentially identical relationships, we collapsed these into a single category of overall tendency toward having thoughts with emotion (of any valence). All results are summarized in Figure 20.
Figure 20. Relationship between individual thought tendencies and mean fractional anisotropy (FA) in white matter tracts throughout the brain. Standardized residual values are displayed (Z-scores) for partial correlations controlling for age and sex. A thumbnail image of each white matter tract is displayed for ease of interpretation. CST: corticospinal tract; CgH: cingulum (hippocampal component); IFO: inferior fronto-occipital fasciculus; tSLF: temporal component of the superior longitudinal fasciculus; Unc: uncinate fasciculus.
Discussion

Grey matter concentration and intentionally directed thought

Some of our most intriguing results were observed for those participants who had higher rates of intentionally directed thought. In previous work, we have speculated that executive frontoparietal control network areas recruited during self-generated thought might be specifically involved in directing, guiding, and selecting among thoughts (Christoff et al., 2009, Fox and Christoff, 2014, Fox et al., 2015, Ellamil et al., 2016, Fox et al., 2016), but to date no empirical evidence has addressed this issue. Here, we provide the first direct evidence that the executive system is indeed preferentially tied to the direction and guidance of self-generated thoughts. This evidence is also consistent with our fMRI results (Chapter 4), which found preliminary evidence that dorsal anterior cingulate cortex, a canonical executive control region, is more active during intentionally directed thought. Here, however, we observed more widespread recruitment of executive areas, including dorsal anterior cingulate cortex, left dorsolateral prefrontal cortex (BA 46), right rostrolateral prefrontal cortex (BA 10/46), and both medial and lateral portions of premotor cortex (BA 6).

Grey matter concentration and goal-related thought

Individual propensity toward goal-related thought was correlated with increased grey matter concentration in a combination of default and frontoparietal control network areas and subcortical regions. Clusters that fell within the boundaries of the frontoparietal control network included the left dorsolateral prefrontal cortex, dorsal anterior cingulate cortex/anterior mid-cingulate cortex, and a small cluster in the left anterior inferior parietal lobule. Clusters were also
observed in default network areas such as the posterior cingulate cortex and right parahippocampal cortex. Finally, significant associations were also found in subcortical regions, including the thalamus and cerebellum.

As with intentionally directed thought, we have previously hypothesized that executive brain regions might be preferentially involved in future planning and other forms of thinking about one’s goals and current concerns (Christoff et al., 2009, Fox and Christoff, 2014, Fox et al., 2015, Ellamil et al., 2016, Fox et al., 2016). The present morphometric neuroimaging results, together with our functional neuroimaging results (Chapter 4), provide the first direct evidence that this is indeed the case. Note that grey matter concentration in these regions appears to contribute independently to both the intentional direction and guidance of thought (previous section), and to a higher propensity to have goal-related thoughts. As we showed above (Chapter 4), intentionally directing thoughts does not necessarily lead them to be more goal-related, and considering the very modest correlation ($R^2 = .014$) between the frequency of these two thought types, for practical purposes they can be considered orthogonal.

**Grey matter concentration and the valence of thought**

Our results for grey matter concentration associated with thought valence were somewhat puzzling, and generally inconsistent with our fMRI results (Chapter 4). Recall that we also found that rates of positive affect during the behavioral testing session were drastically different in the same subjects in a later fMRI scanning session (see Chapter 5), as opposed to the frequency of all other thought qualities, which was extremely stable within individuals and across time-points and contexts (laboratory vs. MRI scanner). We therefore suspect that our ~1 hr thought sampling session was insufficient to capture ‘trait’ affective styles; our reports more likely reflect transient
moods in our participants. We therefore consider it premature to discuss these results further here.

White matter integrity and self-generated thought tendencies

Above, we reported the results of an exploratory pilot study examining relationships between fractional anisotropy (indexed here by mean fractional anisotropy) and individual differences in various thought tendencies. Although this analysis is highly exploratory and data collection is ongoing to improve the power and reliability of our results, a brief discussion is offered here.

One intriguing finding is that mean FA is negatively correlated with rate of stimulus-independent thinking in the corticospinal (i.e., pyramidal) tracts bilaterally. This is intriguing because the corticospinal tracts are the major output pathways from the motor cortex to the body’s voluntary musculature, including controlling fine motor skills (Haines, 2004, Krebs, 2011) – i.e. the main pathways mediating voluntary navigation through the world and manipulation of the body within it. Without speculating unduly, it is interesting to note that individuals who think less about their immediate environment, and who instead engage in higher rates of thought decoupled from the here and now, show reduced fractional anisotropy in the major tracts mediating voluntary control of the body and engagement with the external world.

We also found that the frequency of emotionally-valenced (as opposed to neutral) thoughts was associated with higher FA in the right, but lower FA in the left, medial temporal lobe (hippocampal) components of the cingulum. The cingulum travels beneath the cingulate gyrus and wraps around at its posterior end to travel through and connect with the medial temporal lobe (Haines, 2004, Krebs, 2011). As such, it is the major pathway allowing for
communication between various limbic structures critically involved in emotion, reward, and memory, and has been implicated in affective disorders such as depression as well as disorders characterized by memory impairment, such as Alzheimer’s disease and Mild Cognitive Impairment (Steele et al., 2008, Jones et al., 2013). Although we offer no explanation for the lateralization of these effects, it is intriguing to observe that pathways with an established role in linking limbic brain structures key to emotional processing show anatomical heterogeneities associated with the reported frequency of self-generated emotions.

We also found that the frequency of goal-related thoughts was negatively associated with mean FA in the medial temporal lobe component of the left cingulum, as well as with higher mean FA in the right uncinate fasciculus. The uncinate fasciculus provides major connections between the medial temporal lobe and prefrontal and orbitofrontal cortex (Kier et al., 2004). Consistent with this finding, recall that grey matter concentration was also significantly increased in the right medial temporal lobe (parahippocampal cortex).

Finally, frequency of directed thought was inversely correlated with mean FA in the right corticospinal tract and right inferior fronto-occipital fasciculus, as well as directly correlated with mean FA in the temporal component of the right superior longitudinal fasciculus. With respect to the corticospinal tract, and similar to our comments above, it is interesting to observe decreased fractional anisotropy in the brain’s major motor output tract associated with an increased tendency to control and guide thoughts detached from the external environment. The inferior fronto-occipital fasciculus provides the major pathway conveying visual and auditory information to the frontal lobes; again, it is interesting to note decreased fractional anisotropy in a pathway conveying external sensory information coupled to a stronger tendency to focus executive resources on internal channels of information. The superior longitudinal fasciculus
represents one of the brain’s major rostrocaudal fiber pathways, linking widespread parietal and prefrontal areas, with putative roles in the top-down control of attention and working memory, among other functions (Makris et al., 2005). Again, though of course speculative, increased fractional anisotropy here is reasonable in individuals who have a greater tendency to direct and guide their internal thought streams.

**Cellular foundations of macroscale morphological changes**

The underlying cellular-level changes that putatively give rise to the many gross structural differences reported here remain very poorly understood, particularly in humans. Nonetheless, numerous candidate cellular mechanisms have been identified (Anderson, 2011, Zatorre et al., 2012, Lovden et al., 2013) and are tentatively summarized in Table 24. Although most of these relationships remain highly speculative, some first steps have been taken toward understanding cellular-level changes that might give rise to brain morphology differences identifiable via morphometric neuroimaging. Preliminary evidence from a mouse model (Lerch et al., 2011), for instance, found that volumetric changes as measured by morphometric MRI methods correlated most strongly with the presence of axon growth cones (as opposed to neuron size or number). This suggests that axon sprouting and reorganization may be a more important factor than, for example, neuron soma growth or neurogenesis proper. In contrast, neurogenesis – which results in the addition of relatively few total neurons, remains equivocal in humans, and seems mostly restricted to the hippocampus (Zatorre et al., 2012) – is a very unlikely cellular mechanism (for detailed discussion, see the reviews noted above). Determining the cellular candidates for changes observed with respect to individual differences in patterns of thinking faces the additional challenge that direct animal models are not possible, whereas motor skill or
spatial learning, for instance, can be plausibly simulated in rodents and monkeys and compared concurrently with human results (Sagi et al., 2012).

**Table 24.** Candidate cellular mechanisms putatively underlying macroscale structural differences in gray and white matter

<table>
<thead>
<tr>
<th>Cellular Mechanism</th>
<th>Tissue Type Affected</th>
<th>Morphometric Methods Putatively Affected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dendritic arborization/synaptogenesis</td>
<td>GM</td>
<td>GMC, GMV</td>
</tr>
<tr>
<td>Neurogenesis</td>
<td>GM</td>
<td>CT, GMC, GMV</td>
</tr>
<tr>
<td>Angiogenesis</td>
<td>GM + WM</td>
<td>CT, GMC, GMV, WMV</td>
</tr>
<tr>
<td>Axon sprouting</td>
<td>GM + WM</td>
<td>CT, GMC, GMV, FA</td>
</tr>
<tr>
<td>Gliogenesis</td>
<td>GM</td>
<td>CT, GMC, GMV</td>
</tr>
<tr>
<td>Fiber reorganization</td>
<td>WM</td>
<td>FA</td>
</tr>
<tr>
<td>Myelin formation</td>
<td>WM</td>
<td>FA</td>
</tr>
<tr>
<td>Myelin remodeling</td>
<td>WM</td>
<td>FA</td>
</tr>
<tr>
<td>Astrocyte changes</td>
<td>WM</td>
<td>WMV</td>
</tr>
</tbody>
</table>


**Limitations**

The most significant limitation of these two studies is the uncertainty involved in using thought content measured during a single ~1 hr session as a putative ‘trait’ measure of individual self-generated thought tendencies. Several factors in our design mitigate this concern, however.

First, we collected a large number of thought probes per subject ($n = 120$). To our knowledge, this is far more thought probes than used in any prior neuroimaging study of self-generated thought: Christoff and colleagues (2009), for instance, presented only 16 thought probes per subject; Bernhardt and colleagues (2014), in their morphometric investigation of task-unrelated thought, employed only 7 thought probes (but still used this data as a ‘trait’ measure to
predict brain structure differences in cortical thickness). The number of thoughts sampled here was therefore approximately an order of magnitude larger than in previous work.

Moreover, there is evidence that self-generated thought frequency in the laboratory mirrors thoughts experienced in daily life (McVay et al., 2009), and, as we showed above (Chapter 5), overall self-generated thought tendencies appear to be consistent across both the laboratory and MRI scanner environments. Moreover, within the same participants but across several days, we found that thought content is highly similar, both in and outside the MRI scanner (Chapter 5), further supporting the validity of these individual differences in reported content as putative trait measures.

More typically, ‘trait’ variables are measured with self-report questionnaires. Although some such questionnaires have been developed to capture qualities of self-generated thought (Singer and Antrobus, 1970, 1972), evidently there are serious limitations to this method of identifying ‘trait’ patterns of thought. Such questionnaires employ retrospective self-assessment, which is potentially limited by distortions of memory (Loftus, 1996) and self-serving bias (Shepperd et al., 2008). Moreover, even though a questionnaire might ask the participant to describe average or ‘trait’ qualities, this is no guarantee that temporary state effects (such as current mood or the day’s events) will not color the participant’s responses – the exact same limitation of using moment-to-moment thought sampling as a putative trait measure.
CHAPTER 7 – IDENTIFYING THE NECESSARY AND SUFFICIENT NEURAL BASIS OF SELF-GENERATED THOUGHT: THE VALUE OF NEUROPSYCHOLOGICAL LESION STUDIES

Introduction

All of the aforementioned results are, at heart, correlational: although we have reported many novel relationships between subjective experience, neurophysiological activity, and neuroanatomical structure, nonetheless these associations cannot demonstrate causal roles for any of these brain regions or networks in self-generated thought. In order to provide strong evidence for a causal role of a given brain region in a given cognitive process or mental state, the gold standard is a true loss-of-function experiment in which the area of interest is ablated or inactivated. Although this kind of research is routinely conducted in animal models, human studies must rely on naturally-occurring or surgically-unavoidable lesions. Meticulous observation and assessment of cognitive-affective deficits following circumscribed brain lesions is perhaps the oldest form of ‘cognitive’ neuroscience, dating back to the pioneering work, for instance, of Paul Broca (Broca, 1861). Widespread investigation of the cognitive-affective sequelae of virtually every conceivable form of brain lesion has since developed into the highly influential field known as ‘neuropsychology.’

Neuropsychology studies the cognitive, affective, and perceptual deficits suffered by patients with various brain lesions. Far from merely cataloguing the effects of rare brain disorders and diseases, however, lesion studies have shed much light on healthy brain functioning, as well as serving as highly effective catalysts for further research (e.g., (Scoville and Milner, 1957, Luria, 1976, Solms, 1997, Gainotti, 2000, Müller and Knight, 2006). Lesion studies can provide important clues about the necessary and sufficient neural basis of a given
cognitive process, even a complex one like self-generated thought (Koenigs et al., 2007). To anticipate our conclusions, lesion work suggests that at least four regions play critical roles in various kinds of self-generated thought: medial prefrontal cortex, inferior parietal lobule, medial occipitotemporal cortex, and medial temporal lobe (Fig. 21).

Identifying necessary and sufficient neural substrates of self-generated thought

Although we are aware of little neuropsychological work that has directly addressed the effect of brain lesions on self-generated thought, many studies have addressed closely related cognitive processes. Likely the closest parallel is the extensive neuropsychological lesion work undertaken by Solms regarding the necessary and sufficient neural basis of dreaming (Solms, 1997, 2000b). Solms concluded that the areas most critical to dreaming in general are (i) the medial prefrontal cortex and (ii) the temporoparietal junction/inferior parietal lobule. Additionally, a large swathe of (iii) medial occipital cortex, centering on the lingual gyrus, is critical for the visual features of dreaming (Solms, 1997, 2000b, Bischof and Bassetti, 2004). Elsewhere, we have argued at length that the subjective experiences and neurophysiological correlates of dreaming bear a strong resemblance to those of waking mind-wandering and related forms of self-generated thought (Fox et al., 2013, Fox and Christoff, 2014, Domhoff and Fox, 2015). Notably, all three regions identified by Solms as critical to nighttime dreaming emerged as significantly activated in our meta-analysis of waking self-generated thought (Fox et al., 2015; for further discussion see Domhoff & Fox, 2015). Further supporting this view is the finding that, concurrent with the global loss or reduced frequency of dreaming, patients often reported reduced daydreaming and fantasy following the lesions to medial prefrontal cortex and/or the periventricular white matter tracts at the anterior horns of the lateral ventricles and the genu of
the corpus callosum (Frank, 1946, 1950, Piehler, 1950, Schindler, 1953). To our knowledge, these latter few investigations are the only studies to have directly assessed some form of waking self-generated thought in relation to brain lesions, and unfortunately the results are largely anecdotal. Nevertheless, overall this convergence across methods (neuroimaging and lesion work) suggests that more rigorous investigations of the quality and content of waking self-generated thought in patients with damage to these three areas could prove informative.

Because self-generated thought so often involves spontaneous memory retrieval (Andrews-Hanna et al., 2010, Fox et al., 2013), lesion work related to semantic and autobiographical memory capacity is also relevant. Solms’ conclusions concerning the importance of the medial prefrontal cortex to self-generated waking and dreaming experience are bolstered by work showing that lesions to medial prefrontal cortex have detrimental effects on both semantic and episodic autobiographical memory retrieval (Philippi et al., 2014).

Another potentially critical region, largely disregarded by Solms (1997), is (iv) the medial temporal lobe (Domhoff and Fox, 2015). Bilateral medial temporal lobe lesions are most famously associated with anterograde (and limited retrograde) amnesia (Scoville and Milner, 1957, Milner et al., 1968); less well known are heavy deficits in dreaming that cannot simply be explained by failures of recall. For instance, Korsakoff’s syndrome patients with bilateral medial temporal lobe damage show a marked decrease in dream reports, even when awakened directly from REM sleep in a laboratory setting (Greenberg et al., 1968). Amnesia alone cannot explain these findings, for at least two reasons. First, medial temporal lobe patients still had some dream recall, demonstrating that they still possessed basic recall capacities (e.g., in one study, ~25% of awakenings elicited dream reports (Greenberg et al., 1968), vs. about 80-90% in normal subjects (Hobson et al., 2000). Second, medial temporal lobe patients retain intact working/short-term
memory (Milner et al., 1968): anterograde amnesia should not prevent them from reporting upon experiences from just moments ago, immediately upon being awakened in a sleep laboratory setting. The dreams medial temporal lobe patients do report also support this interpretation: “the content of that material which was recalled showed very stereotyped, commonplace features and reflected very little affect” (Greenberg et al., 1968; p. 205; emphasis added). Nor are these findings restricted to Korsakoff’s patients: a study of encephalitis patients with severe medial temporal lobe lesions reports identical findings (Torda, 1969): patients reported far fewer dreams than controls, and reports “were short and simple… [they] contained one scene with recurrent repetition… The dreams lacked imaginative, unusual, or mysterious details and intensive emotions. The content was stereotyped, repetitious…” (p. 280). The reduction in dream frequency with medial temporal lobe lesions is certainly severe enough to warrant further research with respect to the effect of similar lesions on waking self-generated thought.

Several studies from recent years provide further corroborative evidence for the importance of the medial temporal lobe in self-generated thought processes. For instance, medial temporal lobe patients are severely impaired in imagining novel fictitious or future events and experiences (Klein et al., 2002, Hassabis et al., 2007, Rosenbaum et al., 2009, Andelman et al., 2010, Kwan et al., 2010, Race et al., 2011, 2013), congruent with the consistent activation of medial temporal lobe during such tasks in healthy subjects (Schacter et al., 2007, Addis et al., 2009, Schacter et al., 2012). Medial temporal lobe lesion patients also suffer deficits in tasks requiring creative and novel patterns of thinking (Duff et al., 2013, Rubin et al., 2014). Although providing only circumstantial support, these findings are intriguing because so much of the content of self-generated thought consists of imagined future scenarios and novel, creative recombination of prior memories into new thoughts and simulated experiences (Klinger, 2008,
Andrews-Hanna et al., 2010, Fox et al., 2013). Investigation of the quality and frequency of self-generated thought in medial temporal lobe lesion patients therefore appears to be a promising avenue for future research.
Figure 21. Neuropsychological lesion studies and meta-analysis of fMRI investigations converge on four brain structures that may be critical to self-generated thought. The left column shows brain templates with dark areas indicating regions critical to nighttime self-generated thought (dreaming) as determined by overlapping lesion sites based on CT scans of neurological patients – except panel g, which shows the average site of medial temporal lobe excisions during surgery for intractable epilepsy. In the case of lesions to medial prefrontal cortex (panel a) there is evidence for severe reduction in waking self-generated thought (daydreaming/fantasy) as well. The middle column shows meta-analytic brain activations associated with waking daydreaming/mind-wandering. The right column shows meta-analytic brain activations associated with REM sleep, which is nearly always accompanied by dreaming ($r = .8$; Hobson et al., 2000). Note that the activation cluster in panel f is somewhat anterior to the temporoparietal junction, approximately in Brodmann area 40. All three approaches converge on four areas: the medial prefrontal cortex (panels a–c); the temporoparietal junction/inferior parietal lobule (panels d–f); the medial temporal lobe (panels g–i); and the medial occipital lobes/lingual gyrus (panels j–l). Panels a and d based on the work of Solms (2000a); panel j based on Solms (1997). Panels b, e, and h based on data from Fox et al. (2015). Panel g based on average medial temporal lobe excisions for severe epilepsy patients as reviewed by Mathern and Miller (2013). Panels c, f, and i based on data from Fox et al. (2013) and Domhoff & Fox (2015). X and Z values represent left-right and vertical coordinates, respectively, in Montreal Neurological Institute stereotactic space. Figure expanded and modified from Domhoff & Fox (2015), with permission.
CHAPTER 8 – NEURAL ORIGINS AND ONTOGENY OF SELF-GENERATED THOUGHT AS REVEALED BY HUMAN INTRACRANIAL ELECTROPHYSIOLOGY

Introduction

Neither functional neuroimaging nor lesion studies can answer questions about the detailed, millisecond-scale temporal dynamics of self-generated thought. Two key questions about these temporal dynamics concern the neuroanatomical origins and ontogeny of self-generated thought: where in the brain do self-generated thoughts tend to originate (there may of course be more than one answer), and how does self-generated activity subsequently spread through distributed neuronal networks and give rise to the accompanying subjective experiences of memory recall and novel thought? A third important question centers on the dynamics of interactions between large-scale networks, most importantly the frontoparietal control and default mode networks, which we have discussed in detail elsewhere (Christoff et al., in press).

Similar to the lesion studies discussed above (Chapter 7), the ideal method for investigating fine timescale neural activity – intracranial electrophysiology – is only permissible in humans in clinical contexts (most commonly, epilepsy patients). While we therefore did not conduct such an investigation here, much can be gained from a careful review of the many scattered reports of human electrophysiological recording and stimulation experiments over the past century. In this chapter we review the results of more than 100 such studies, expertly synthesized in the comprehensive review of Selimbeyoglu and Parvizi (2010), and summarize what has been learned about brain regions key to the origins (initial generation) and ontogeny (subsequent propagation) of self-generated forms of thought.
Neural origins of self-generated thought

With respect to sites of origin, there may be more than one central location of generation, or it may be that difficult-to-localize, distributed network activity gives rise to thoughts, and the search for ‘primary’ thought generation regions is misguided. The reality could also be that both mechanisms contribute, depending on the type and content of self-generated thought. Whatever the answers to these quandaries, the evidence for the central role of the medial temporal lobe in thought generation is compelling. The most direct causal evidence comes from the ever-growing body of cognitive studies of patients who have had brain electrodes chronically implanted, or the cortical surface probed with electrical stimulation, for a variety of clinical reasons – usually intractable epilepsy. Since the first studies of this kind nearly one hundred years ago, essentially every brain area has been explored with direct focal electrode stimulation, or passive recording of spontaneous discharges, to a greater or lesser degree. The most relevant finding from this growing body of research is that stimulation of medial temporal lobe structures (i.e., the hippocampus, parahippocampus, entorhinal cortex, and amygdala) very frequently leads to memory recall, immersive thoughts, and hallucinatory, dream-like experiences (Fig. 22). Far from an isolated occurrence, the accumulated evidence supporting this assertion is fairly substantial (Feindel and Penfield, 1954, Bickford et al., 1958, Baldwin, 1960, Penfield and Perot, 1963, Horowitz et al., 1968, Ferguson et al., 1969, Halgren et al., 1978, Wieser, 1979, Gloor et al., 1982, Fish et al., 1993, Bancaud et al., 1994, Kahane et al., 2003, Barbeau et al., 2005, Vignal et al., 2007, Mulak et al., 2008, Jacobs et al., 2012).

Equally important, however, is the specificity of these results: a comprehensive review of cognitive-affective findings from electrophysiological studies in humans, summarizing the results of over 100 such studies conducted over the past eighty years, found that such phenomena
were produced almost exclusively by stimulation of medial temporal lobe structures (Selimbeyoglu & Parvizi, 2010; see our summary of relevant results in our Table 25). Stimulation of nearby temporopolar cortex (Penfield and Perot, 1963, Halgren et al., 1978, Bancaud et al., 1994) or lateral temporal cortex (Penfield, 1958, Mullan and Penfield, 1959, Penfield and Perot, 1963, Bancaud et al., 1994) also occasionally elicits such phenomena, but these reports are comparatively rare, and moreover electrical stimulation is known to spread to adjacent cortical areas, making it difficult to strictly rule out a medial temporal origin (or co-activation) in many of these cases (Gloor et al., 1982, Gloor, 1990, Bancaud et al., 1994). Indeed, many of the stimulations to lateral temporal cortex that elicited thought- and dream-like experiences were reported from electrodes at a depth of several centimeters, supporting such a possibility (Selimbeyoglu and Parvizi, 2010). Stimulation of lateral prefrontal cortex (Blanke et al., 2000a) and orbitofrontal cortex (Mahl et al., 1964) can likewise occasionally elicit such experiences, but such reports are few and far between (Table 25). Most striking is that no such experiences appear to have ever been reported from stimulation of virtually any other area in the brain (Selimbeyoglu and Parvizi, 2010), even from the many other regions consistently recruited by self-generated thought in functional neuroimaging investigations. For instance, fewer than 1% of stimulations to the inferior parietal lobule elicit such phenomena (Selimbeyoglu and Parvizi, 2010). We summarize these results in Table 25 and Fig. 22.

Although absence of evidence is not necessarily evidence of absence, the results are intriguing given the number of investigations already carried out and the span of time over which such investigations have been taking place (well over a hundred studies, over nearly a hundred years). Together these results strongly suggest that the medial temporal lobe is a key generation
site for many forms of self-generated thought, including dreaming – a hypothesis that could be examined more forcefully with lesion patients (see previous section).

A central limitation of drawing conclusions based on this body of research is that the results were primarily evoked by electrical stimulation – they mostly do not represent self-generated, spontaneous brain activity giving rise to immersive thoughts and hallucinatory, dream-like experiences. Several studies, however, have passively recorded spontaneous brain activity in conjunction with first-person reports of accompanying experience, and yielded consistent findings: the medial temporal lobe appears to be by far the most common origin of self-generated brain activity accompanied by spontaneous memories, thoughts, and dream-like experiences (Bancaud et al., 1994, Vignal et al., 2007, Gelbard-Sagiv et al., 2008).
Figure 22. Preferential involvement of medial temporal lobe structures and temporopolar cortex in electrophysiological stimulations (or spontaneous discharges) eliciting memories, thoughts, or hallucinatory, dream-like experiences. Percentage of stimulations or spontaneous discharges that elicited a first-person experience of memories, thoughts, or hallucinatory, dream-like experiences, based on more than 100 independent investigations. Not shown are data for hundreds of other stimulations throughout the brain, for which no such thought- or dream-like experiences have ever been reported. Only brain areas with ≥10 stimulations or discharges reported in the literature are visualized. Drawn from data in our Table 25, based on data in Supplementary Table 1 of the comprehensive review of Selimbeyoglu & Parvizi (2010). HPC: hippocampus; IPL: inferior parietal lobule; ITG: inferior temporal gyrus; MTG: middle temporal gyrus; PHC: parahippocampal cortex; STG: superior temporal gyrus; TOJ: temporo-occipital junction; TPC: temporopolar cortex.
**Table 25.** Summary of human electrophysiology studies demonstrating elicitation of memories, thoughts, or hallucinatory, dream-like experiences.

<table>
<thead>
<tr>
<th>Brain Region</th>
<th>Stimulations/discharges eliciting</th>
<th>Total stimulations/discharges</th>
<th>Percentage eliciting</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Temporal Lobe</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hippocampus</td>
<td>25</td>
<td>46</td>
<td>54%</td>
<td>(Halgren et al., 1978, Fish et al., 1993, Bancaud et al., 1994, Kahane et al., 2003, Vignal et al., 2007, Mulak et al., 2008)</td>
</tr>
<tr>
<td>Amygdala</td>
<td>13</td>
<td>36</td>
<td>36%</td>
<td>(Ferguson et al., 1969, Halgren et al., 1978, Fish et al., 1993, Vignal et al., 2007)</td>
</tr>
<tr>
<td>Parahippocampal region</td>
<td>9</td>
<td>16</td>
<td>56%</td>
<td>(Feindel and Penfield, 1954, Penfield and Perot, 1963, Vignal et al., 2007)</td>
</tr>
<tr>
<td>Temporopolar cortex</td>
<td>5</td>
<td>11</td>
<td>45%</td>
<td>(Penfield and Perot, 1963, Halgren et al., 1978, Bancaud et al., 1994, Ostrowsky et al., 2002, Mulak et al., 2008)</td>
</tr>
<tr>
<td>Inferior temporal gyrus</td>
<td>1</td>
<td>21</td>
<td>5%</td>
<td>(Penfield and Perot, 1963)</td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>7</td>
<td>42</td>
<td>17%</td>
<td>(Penfield, 1958, Mullan and Penfield, 1959, Penfield and Perot, 1963, Kahane et al., 2003)</td>
</tr>
<tr>
<td>Temporo-occipital junction</td>
<td>4</td>
<td>17</td>
<td>24%</td>
<td>(Penfield and Perot, 1963, Morris et al., 1984, Lee et al., 2000)</td>
</tr>
<tr>
<td><strong>Frontal Lobe</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>1</td>
<td>7</td>
<td>14%</td>
<td>(Blanke et al., 2000a)</td>
</tr>
<tr>
<td>Middle frontal gyrus</td>
<td>2</td>
<td>8</td>
<td>25%</td>
<td>(Blanke et al., 2000a)</td>
</tr>
<tr>
<td>Orbitofrontal cortex</td>
<td>1</td>
<td>4</td>
<td>25%</td>
<td>(Mahl et al., 1964)</td>
</tr>
<tr>
<td>Supplementary motor area</td>
<td>1</td>
<td>6</td>
<td>17%</td>
<td>(Beauvais et al., 2005)</td>
</tr>
<tr>
<td><strong>Parietal Lobe</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inferior parietal lobule</td>
<td>2</td>
<td>42</td>
<td>0.05%</td>
<td>(Blanke et al., 2000b, Schulz et al., 2007)</td>
</tr>
</tbody>
</table>

Based on data in Supplementary Table 1 in the comprehensive review conducted by Selimbeyoglu & Parvizi (2010). Data for brain areas with ≥10 stimulations/discharges reported in the literature are visualized in Fig. 22.
The subsequent ontogeny of self-generated thought

Even highly localized stimulation or spontaneous neural firing will tend to spread to proximate and distant brain areas through short- and long-range connections (Halgren and Chauvel, 1992, Selimbeyoglu and Parvizi, 2010). Assuming that the medial temporal lobe is one of the main sources of spontaneous activity giving rise to thoughts and spontaneously recalled memories, the question remains of how and on what timescale this activity spreads, and where in the brain it is most likely to spread to. Given that the medial temporal lobe is well connected to virtually every other part of the brain (Buzsáki, 1989, Rolls, 2000, Simons and Spiers, 2003, Buzsaki, 2006), the possibilities are legion.

Our view of the role of the medial temporal lobe in spontaneously reactivating memory traces, and in recombining mnemonic material into novel thoughts and imaginings, draws heavily on the hippocampal indexing theory (Teyler and DiScenna, 1985, 1986, Moscovitch, 1992, Teyler and Rudy, 2007). In essence, “indexing theory proposes that the content of our experiences is stored in the multiple neocortical loci activated by experience and the hippocampus stores an index of those neocortical loci” (Teyler & Rudy, 2007; p. 1160). Indexing theory is primarily concerned with accounting for the reactivation of memories; our view expands on this to account also for novel patterns of thought that draw on mnemonic material but recombine it into new thoughts and simulations of experience. If a specific pattern of medial temporal lobe activity and connectivity to neocortex indeed gives rise to memory recall, it follows that slightly altered or novel patterns of hippocampal activity will ‘index’ novel patterns of distributed neocortical activity never before experienced – that is, novel thoughts and simulated experiences. The central problem is to understand “the manner in which electrical stimulation of a particular location [or spontaneous, self-generated activity at a given location]
leads to activation of [a] very widespread but, at the same time, very particular network” (Bancaud et al., 1994; p. 87).

We therefore do not predict a single, universal pattern of spreading activity. The pattern of subsequent recruitment should instead be related to the experiential qualities of the accompanying thoughts, including the sensory modalities instantiated (visual, auditory, somatosensory, etc.); affective tone (positive, negative, or neutral); and other qualities such as temporal orientation (past, present, future) and goal-relatedness. For instance, thoughts that are visual in nature may involve spreading activity (at the neuronal level) and subsequently observable recruitment (e.g., with functional neuroimaging) of areas such as the lingual gyrus. Similarly, we would predict that more goal-related thoughts involving planning for the future should tend to recruit prefrontal executive areas such as dorsal anterior cingulate cortex and rostrolateral prefrontal cortex.

A further consideration is that new thoughts and imagined mental content may not be determined merely by patterns of synaptic connections alone – these anatomical connections could also interact with a host of other factors that all participate in the sculpting of spontaneous and ‘noisy’ brain activity into an overall pattern of activity ultimately corresponding to subjectively experienced mental content of one form or another. Aside from the basic neurochemical state of the brain at a given time (see next section), “current sensory input, cognitive context and/or psychosocial concerns could sometimes be major influences in defining the final pattern to emerge from the sculpting” (Bancaud et al., 1994; p. 87).

Evidence bearing on this question is unfortunately much more sparse – commensurate with the increased difficulty of the problem. A comprehensive understanding of the spread of spontaneous activations would require widespread electrode placement throughout the brain – a
situation uncommon in human electrophysiology studies, where electrode placement is
determined strictly by clinical criteria and rarely requires simultaneous recording from many
widely dispersed sites.

To truly explore the hypothesized model of self-generated thought ontogeny requires
recording from relevant cortical areas and the medial temporal lobe, simultaneously. The study
that probably best approximates this ideal methodology investigated episodic memory retrieval
while recording simultaneously from medial temporal lobe and the retrosplenial cortex in the
posteromedial area (Foster et al., 2013). The authors reported that episodic memory retrieval
involved phase locking in the theta band (3-4 Hz) between these two regions, and that this
concerted activity was unique to this frequency band as well as to these two regions. Most
intriguing is that the coupling was strongest prior to actual peak high-frequency activity in the
retrosplenial cortex, potentially suggesting a primary role for the medial temporal lobe. Another
study investigating spontaneous memory recall with human intracranial electrophysiology found
that medial temporal lobe structures demonstrated some of the strongest increases in high-
frequency (i.e., high γ-band) activity just prior to conscious recall and reporting of the memory
(Burke et al., 2014), whereas high-frequency activity peaked later in several parietal, temporal,
and frontal regions (see their Fig. 4). Importantly, although the medial temporal lobe was not the
only region in which high-frequency activity peaked prior to recall, it was the only one in which
this activity significantly predicted subsequent memory recall, suggesting a cardinal role in
memory recollection (Burke et al., 2014).

A limitation of these studies (for our purposes) is that they involve memory retrieval in
one form or another; therefore they do not speak to the possibility of truly novel patterns of
activity – a precondition for any neuron-level model of self-generated thought. Some preliminary
evidence for this possibility, however, comes from animal models. Several well-known animal studies have demonstrated that patterns of medial temporal lobe activity reflecting recent spatiotemporal experiences and memories are spontaneously replayed during periods of resting wakefulness (Foster and Wilson, 2006, Diba and Buzsáki, 2007) or subsequent sleep (Wilson and McNaughton, 1994). Coordinated replay across medial temporal lobe and various neocortical areas, including visual regions (Ji and Wilson, 2007) and posterior parietal cortex (Qin et al., 1997), has also been reported – although a causal role for the medial temporal lobe in initiating this activity has not been shown. Interestingly, recent studies have shown that ‘replay’ need not be a mere recapitulation of previous firing patterns, but often contains novel firing sequences that do not correspond to any actual spatiotemporal sequence of experience (Davidson et al., 2009, Gupta et al., 2010). These novel firing sequences have been interpreted as evidence for planning and imagining of alternative behaviors (Knierim, 2009, Gupta et al., 2010) – i.e., the generation of novel thoughts, plans, and imagined scenarios, akin to self-generated thought content in humans.

In summary, although the network-level ontogeny of thoughts putatively generated in the medial temporal lobe remains largely obscure, many important prerequisites for an ‘indexing’ scenario have already been demonstrated. The medial temporal lobe is known to spontaneously reactivate patterns of activity first instantiated during novel experience and learning, and this reactivation takes place not only during sleep but also waking behavior and restful states. This patterned replay can be temporally synchronized with various neocortical brain areas (Qin et al., 1997, Ji and Wilson, 2007), and, critically, ‘replay’ can in fact involve novel patterns of activity that do not correspond to any specific experience (Davidson et al., 2009, Knierim, 2009, Gupta et al., 2010). Although still far from definitive, all of these findings are consistent with an
‘indexing’ theory of self-generated thought origin and ontogeny, whereby spontaneous medial temporal lobe activity activates neuronal networks dispersed throughout the brain, instantiating either recall of memories or the experience of novel thoughts and imaginings. Perhaps most important, electrode montages that include both medial temporal lobe and other regions clearly important to self-generated thought are occasionally employed in clinical settings with human participants (Ekstrom et al., 2003, Gelbard-Sagiv et al., 2008, Foster et al., 2013, Burke et al., 2014); the stage is therefore set for further exploration of the temporal and spatial dynamics that characterize the ontogeny of thoughts in the human brain. The most fruitful approach might be to combine intracranial electrophysiology (which has unparalleled temporal resolution, but cannot be expected to cover a large number of brain regions in human studies) with functional MRI (with poor temporal resolution, but the ability to sample from all brain regions simultaneously). The safe and optimal combination of these methods is an area of active research and has already been demonstrated in feasibility studies (Carmichael et al., 2007, Carmichael et al., 2010).
CHAPTER 9 – SUMMARY AND CONCLUSIONS

In this final chapter we briefly summarize the key findings from each of the seven preceding chapters and attempt to integrate them to draw some general conclusions about the neural basis of self-generated thought.

Self-generated thought in general is associated with widespread brain recruitment

Two central conclusions can be drawn from our meta-analytic assessment of general brain recruitment associated with various forms of self-generated thought (Experiment 1). First, as was long suspected, the default network participates critically in self-generated thought: every major default network region shows consistent recruitment. However, this is not the whole story: numerous other brain areas and networks also show consistent recruitment, most notably frontoparietal control network areas, including dorsal anterior cingulate cortex and lateral prefrontal regions. Other areas beyond both these networks are also activated consistently, including medial occipital areas, temporopolar cortex, and the insula. This wide variety of brain recruitment accords well with the subjective diversity of self-generated thought, but cries out for more detailed study to resolve the specific role played by these various regions and networks – a major aim of our subsequent Experiments 3, 5, and 6.
Self-generated thought and imagery in sleep and dreaming share a neural substrate with waking thought

Despite the apparent quiescence of sleep, it is now established that all stages of sleep are accompanied to some degree by self-generated thought and imagery (Nielsen, 2000). The relative frequency of this sleep mentation differs markedly between stages, peaking in REM sleep and bottoming out in SWS. Our meta-analyses of all functional neuroimaging investigations that have explored the various stages of sleep revealed intriguing correspondences between the relative frequency of such self-generated thought and imagery, and activation in default and visual network areas, as expected based on our earlier work along these lines (Fox et al., 2013, Domhoff and Fox, 2015). We also found a similar correspondence between deactivation of frontoparietal control networks regions and the corresponding loss of cognitive control and metacognitive awareness in most sleep stages (with the notable exception of lucid REM sleep, during which activation in these regions is restored). Aside from arguing forcefully that self-generated thought shares a common neural basis regardless of conscious state (sleeping or awake), these results show that self-generated thought can easily take place without executive system recruitment, therefore setting the stage for our subsequent experiments that aimed to specifically investigate the role played by executive regions (Experiments 3 and 5).

Distinct categories of thought can be resolved at the neural level with fMRI

Our fMRI results point to the ability of functional neuroimaging to resolve specific categories of thought at the neural level, based merely on large numbers of first-person reports. Although we concluded that, despite our large number of thought probes (*n* = 120), our analyses were still under-powered, nonetheless important distinctions were drawn between basic neural
correlates for several major categories of thought. The most notable findings, we believe, are the first direct evidence for brain recruitment during self-generated thought involving executive processes, namely intentionally directed/guided thought, and thought related to one’s goals and current concerns. Functional neuroimaging revealed executive system recruitment directly associated with these types of thought (as well as other brain recruitment) – providing the first tentative answer to questions about what role executive regions might be playing in apparently spontaneous and undirected mental processes at rest. These results are further bolstered by our subsequent morphometric neuroimaging analyses, which found concordant results.

**The MRI scanner environment does not appreciably affect self-generated thought content**

The central conclusion of Experiment 4 (Chapter 5) is that the MRI scanner environment does not appear to have any appreciable influence on the overall content or valence of self-generated thought. The rate of stimulus-independent thought might be somewhat higher in the MRI scanner compared to a laboratory setting, but this observation might also have been due to chance differences in stimulus-independent thought frequency between our two samples. These findings apply at the group level (i.e., average thought content does not differ systematically across environmental contexts). A further conclusion, however, can be drawn at the individual level, based on data from our participants who completed the thought sampling paradigm both in and outside of the MRI scanner. These participants, too, showed overall highly stable patterns of thought both in and outside the MRI scanner, and across several days. These findings broadly suggest that a ~1 hr thought sampling session involving >100 thought probes can provide reasonably stable, trait-like measures of individual thought propensity that could
conceivably be used as individual difference measures in further analyses – a task we took up in our subsequent Experiments 5 and 6.

**Interindividual differences in thought content are associated with individual neuroanatomical heterogeneity**

Converting our large number of thought probes into individual difference measures, we examined the possibility of relationships between individual neuroanatomical heterogeneities and subject-specific patterns of thought content. We found many such relationships using measures of both grey matter concentration as well as white matter fractional anisotropy. The most intriguing results from these analyses again pointed to the specific role played by executive brain regions in intentionally directing or guiding the flow of self-generated thoughts, and independently, in facilitating thought related to one’s goals and current concerns. Given our relatively small samples in these correlational analyses, however, and our ongoing collection of participants, these results should be interpreted with some caution for the time being.

**Lesions studies point to necessary neural substrates of self-generated thought**

A synthesis of nearly a century of neuropsychological lesion studies provided telling evidence that only certain regions are absolutely *critical* to self-generated thought. Reviewing this literature, we found that self-generated thought is abolished, or severely attenuated, with lesions to only three brain regions: the medial prefrontal cortex, the inferior parietal lobule/temporoparietal junction, and the medial temporal lobe. Additionally, medial occipital lesions abolish the visual imagery so characteristic of self-generated thought in both waking and
sleeping (see Chapters 1 and 3). The selectivity of these findings is bolstered by the complementary discovery that lesions to many other brain regions have no appreciable effect on self-generated thought frequency or content.

**Human intracranial electrophysiology suggests a medial temporal lobe origin site for self-generated thought**

The preceding work addressed many functional neural qualities of self-generated thought, but none of it addressed perhaps the most intriguing question of all: where in the brain our self-generated mental content originates. Synthesizing nearly one hundred years of isolated human intracranial electrophysiology reports, we concluded that there is strong evidence that the medial temporal lobe is the major (if not necessarily the only) origin site of most self-generated thought. Very little research, however, has addressed the ways in which this self-generated medial temporal lobe activity then propagates to, and recruits, other brain regions – a major avenue for future research.
REFERENCES


American Journal of Psychiatry 152.


cognitive science. Behavioral and Brain Sciences 36:181-204.


Reviews Neuroscience 10.

Cumming G (2012) Understanding the new statistics: Effect sizes, confidence intervals, and
meta-analysis: Routledge.

0956797613504966.

D’Argembeau A, Collette F, Van der Linden M, Laureys S, Del Fiore G, Degueldre C, Luxen A,


Delamillieure P, Doucet G, Mazoyer B, Turbelin M-R, Delcroix N, Mellet E, Zago L, Crivello F,
questionnaire for evaluation of inner experience during the conscious resting state.

Dement W, Kleitman N (1957a) Cyclic variations in EEG during sleep and their relation to
eye movements, body motility, and dreaming. Electroencephalography and clinical
neurophysiology 9:673-690.


Hume RE (1921) The thirteen principal Upanishads: Translated from the sanskrit with an outline of the philosophy of the Upanishads and an annotated bibliography: H. Milford, Oxford University Press.


Krebs C (2011) Neuroscience: Lippincott Williams & Wilkins.


Lilly JC (1956) Mental effects of reduction of ordinary levels of physical stimuli on intact, healthy persons. Psychiatric Research Reports.


Qin Y-L, Mcnaughton BL, Skaggs WE, Barnes CA (1997) Memory reprocessing in
corticocortical and hippocampocortical neuronal ensembles. Philosophical
Transactions of the Royal Society B: Biological Sciences 352:1525-1533.
Race E, Keane MM, Verfaellie M (2011) Medial temporal lobe damage causes deficits in
episodic memory and episodic future thinking not attributable to deficits in
Radua J, Mataix-Cols D (2009) Voxel-wise meta-analysis of grey matter changes in
Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default
mode of brain function. Proceedings of the National Academy of Sciences USA
98:678-682.
Ramón y Cajal S (1928) Degeneration & regeneration of the nervous system: Oxford
University Press, Humphrey Milford.
Rechtschaffen A, Kales A (1968) A manual of standardized terminology, techniques and
scoring system for sleep stages of human subjects.
10:380-388.


