Dynamic Regulation of Internal Experience: Mechanisms of Therapeutic Change

Jessica R. Andrews-Hanna\textsuperscript{1,2}, Kalina Christoff\textsuperscript{3,4}, & Mary-Frances O’Connor\textsuperscript{1}

\textsuperscript{1}Department of Psychology, University of Arizona, Tucson, AZ, USA
\textsuperscript{2}Cognitive Science Program, University of Arizona, Tucson, AZ, USA
\textsuperscript{3}Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada
\textsuperscript{4}Centre for Brain Health, University of British Columbia Vancouver, British Columbia, Canada

\textit{Address correspondence to:}

Jessica Andrews-Hanna
Department of Psychology
1503 E. University Blvd.
Tucson, AZ 85721
jandrewshanna@gmail.com

Citation:
Dynamic Regulation of Internal Experience: Mechanisms of Therapeutic Change

Jessica R. Andrews-Hanna\textsuperscript{1,2}, Kalina Christoff\textsuperscript{3,4}, & Mary-Frances O’Connor\textsuperscript{1}

\textsuperscript{1}Department of Psychology, University of Arizona, Tucson, AZ, USA
\textsuperscript{2}Cognitive Science Program, University of Arizona, Tucson, AZ, USA
\textsuperscript{3}Department of Psychology, University of British Columbia, Vancouver, British Columbia, Canada
\textsuperscript{4}Centre for Brain Health, University of British Columbia Vancouver, British Columbia, Canada

Address correspondence to:

Jessica Andrews-Hanna
Department of Psychology
1503 E. University Blvd.
Tucson, AZ 85721
jandrewshanna@gmail.com

Citation:
KEYWORDS
Default, Thought, Rumination, Depression, Anxiety, Spontaneous, Mindfulness, Memory, Autobiographical, Self

ABSTRACT
Psychotherapy holds great promise for bringing about therapeutic change, yet a major challenge lies in translating short-term change into change that endures over extended time scales. In this chapter, we tackle the “how’s” of therapeutic change through the lens of an emerging field of research on internally-guided experience, encapsulating thoughts and feelings. We first synthesize basic science and clinical research on functional and dysfunctional internal thought, highlighting the importance of alterations in content, processes, and corresponding patterns of functional activity and connectivity of the brain’s default network. Next, we introduce a neurocognitive model highlighting spontaneous processes, deliberate processes, and automatic affective processes that promote or inhibit the dynamics of thought. We apply this dynamic framework to understanding mechanisms of change associated with common psychotherapies, and review preliminary effects of therapy on brain activity and connectivity within and between large-scale brain networks. Throughout the chapter, we note many points of convergence with Lane and colleagues’ “Integrated Memory Model” (IMM), which inspired the edited volume to which this chapter belongs.
1. INTRODUCTION

Recent statistics highlight that approximately one out of every five individuals in the U.S. will experience a mental health disorder in their lifetime\(^1\), and one out of every two will experience symptoms at a minimum of subclinical levels\(^1,2\). The somber reality that nearly half of the population desires improved mental well-being signifies a significant mental health crisis, and ignites an urgency to find solutions that work and are accessible to all who need them. Fortunately, decades of clinical practice have yielded a variety of psychotherapy treatments that promote beneficial change. A major challenge, however, lies in translating short-term change into change that endures over extended time scales\(^3\).

Here we consider the "how's" of change through the lens of an emerging field of research on internally-guided thought and emotion. Despite behaviorist biases confronting basic science research on this topic\(^4\), recent years have brought growing interest in understanding the functional and dysfunctional neurocognitive mechanisms underlying internal experience\(^5-7\). In this chapter, we synthesize basic science and clinical research on internal thought and emotion to consider how to best promote their adaptive outcomes and minimize their maladaptive ones. We first discuss the psychological literature on the normative characteristics, affective content, and proposed functions of internally-guided thought, and then discuss the ways in which it manifests in a dysfunctional manner, signaling compromised well-being. We incorporate neuroscientific perspectives, considering recent research exploring the function and organization of large-scale brain networks in healthy and clinical populations. Finally, we consider how different clinical therapies may modify aspects of thought and emotion, and suggest avenues for future research.

At the core of our framework lies the dynamics of internal experience, as emphasized in our recent neurocognitive model outlining processes constraining the way thoughts and emotions arise and unfold over time\(^6\). We propose that beneficial change over longer timescales may best be understood and promoted by further insight into the dynamics of thinking and underlying brain systems on much shorter time scales.

Throughout the chapter, we note many points of convergence with Lane and colleagues’ “Integrated Memory Model\(^3\)” (IMM), which inspired the edited volume to which this chapter
belongs. The IMM holds that different therapies promote change through different entry points into an overarching system whereby episodic memories, semantic structures, and emotional responses interact to shape internal experience. According to the IMM, and the associated LRNG model, all effective therapies involve updating previously-learned, maladaptive internal experiences with new adaptive experiences introduced as part of the therapeutic practice. With repeated experience associating such adaptive experiences in conjunction with old maladaptive ones across a broad range of contexts, these experiences become transformed over time, increasing the likelihood for beneficial enduring change.

2. INTERNAL EXPERIENCE - WHAT'S NORMAL?

Although a great deal of insight regarding the nature of functional and dysfunctional thinking can be gleaned from introspective techniques, introspection was dismissed as an unsuitable behavioral paradigm or topic of study throughout the behaviorist era of psychology research. Unfortunately, this view has persisted throughout much of modern cognitive science, psychology, and neuroscience. A major consequence of this historical bias is that, despite being so personally familiar to us, relatively little is known about internal experience from a basic science perspective, as thoughts and feelings are otherwise challenging to measure. Yet significant progress towards understanding the precise ways in which these processes can go awry necessitates scientific understanding of internal experience in healthy individuals. This dilemma is even more pertinent when considering mechanisms that may predispose, precipitate, or even ameliorate dysfunctional internal processes.

In 2010, a pair of researchers published an observation that sparked a paradigm shift towards a broader appreciation of the importance of introspective inquiry. Echoing elements of earlier work that had largely been ignored (reviewed in), these researchers demonstrated that when more than 2,000 individuals were surveyed at random moments during everyday life, they were nearly as likely to be thinking about something other than the "task" or "activity" at hand as they were to be thinking about the task itself. The implications for human cognition, and for the way human cognition had traditionally been studied, were far-reaching. The picture painted above implies that adults are drawn to a rich inner mental life often decoupled from external task
demands. The elusive nature of this off-task focus may be mistaken as "noise," yet its ubiquitous nature and phenomenological qualities suggest it has adaptive functions worthy of study, and of relevance to mental health\textsuperscript{11–13}.

\textit{2a. Adaptive Functions of Everyday Thought}

Despite the historical bias alluded to above, a growing number of studies have begun to yield important insight into the phenomenological qualities of off-task and daily thoughts in typical, healthy adults (reviewed in \textsuperscript{5,11,14}). These studies reveal that thoughts tend to concern personally-significant topics, consistent with functional hypotheses that they serve to remind us of, and help us solve, important unresolved "current concerns"\textsuperscript{15}. One aspect of life we often deem personally significant is our social relationships. Accordingly, everyday thoughts sometimes take on a social orientation\textsuperscript{16–18}, perhaps allowing us to make meaning from our interactions with others, predict others' future actions and mental states so we may act accordingly, and incorporate people for whom we care into our sense of self-identity.

In normative samples, a number of studies have demonstrated that typical thought patterns lack strong emotional salience, and are, on average, mildly enjoyable\textsuperscript{19}. Important trial-related variability accompanies these means, with future-oriented thoughts often rated as more positive than past-oriented thoughts\textsuperscript{16}. Mildly positive content may serve to facilitate mental exploration and flexibility, which may promote formation of novel associations and generation of creative solutions to ongoing problems. Indeed, incubations that promote mind-wandering with mindless tasks may boost creative and prospective problem solving\textsuperscript{20}.

Another significant aspect of our lives, correspondingly reflected in our idle moments, is what the future might hold for us. As such, a number of studies have documented a prospective bias to off-task thoughts, including upcoming events upon which we place value, and personal goals we wish to move towards\textsuperscript{21–23}. These findings suggest that a major function of internally-guided thought may be to help us predict, plan, and prepare for what may lie ahead, and act in ways that facilitates movement toward our immediate and long-term goals\textsuperscript{24}. By mentally simulating hypothetical futures and corresponding mental states each may bring out, we may be in a better
position to choose the most optimal route forward. At the same time, internally-guided thoughts may have important mnemonic functions, helping us to consolidate significant past events into long-term memory\textsuperscript{24,25}. Indeed, periods of rest following encoding appear to be beneficial for learning\textsuperscript{26}, and these mental breaks both encourage the emergence of spontaneous thoughts (see below), and functional integration of the medial temporal lobe with cortical regions involved in memory formation\textsuperscript{6}. The mnemonic functions of spontaneous thoughts may also have implications for the transformation of memories over time, a process tied to re-consolidation, and thus relevant for the IMM and LRNG models of Lane and colleagues\textsuperscript{3}.

Paralleling the positive affective content of everyday thought, daily experience sampling paradigms in large groups of individuals have also revealed a prevalence of positive compared to negative emotions, the most frequent positive emotions being joy and love, followed by satisfaction\textsuperscript{27}. Mixed positive-negative emotions are also common, representing nearly one third of sampled occasions. Naturally, the question arises as to whether affective states play a causal role in subsequent thought content, and vice versa. Although studies addressing these questions are relatively scarce, findings from negative mood induction and rumination paradigms suggest that negative affective states increase both the frequency of internally-guided thoughts, and promote more negative, past-focused content, with corresponding physiological correlates. Such findings have been extended to longitudinal daily experience sampling studies, which have additionally revealed that negatively-valenced off-task thought increases the likelihood of subsequent negative mood\textsuperscript{28}.

In summary, a growing number of studies suggest that the content of off-task and daily thoughts of typical, healthy adults is overall positive and constructive, suggesting numerous adaptive functional consequences ranging from problem solving to prediction, and that these thoughts might be at least somewhat related to concurrently-experienced emotions\textsuperscript{10,11}. However, it should be noted that despite the adaptive functions implied by the phenomenological qualities of such thoughts, relatively little research has experimentally quantified their functional consequences. Longitudinal studies in which the antecedents and consequences of different types of thought patterns and emotions on behavior and well-being mark an important direction of future research.
Historical biases and experimental challenges thwarting our understanding of the psychology of internally-guided thought have also limited our understanding of its neural underpinnings. The brain system most closely linked to internal thought was discovered only two decades ago (and entirely by accident, for that matter), as a set of regions that become more engaged during wakeful passive states than during a variety of external tasks\(^29\). These regions would later be coined the "default mode of brain function" by Raichle and colleagues\(^30\), and subsequently the "default mode network"\(^31\) (or "default network"\(^32\) (DN) for short) based on its functional network properties (Figure 1A).

Although the precise contributions of the DN to internal experience remain unclear, a growing number of studies suggest that the DN is better characterized by the internal mental processes it supports than by its opposition to goal-directed tasks (reviewed in \(^14\)). By this account, the DN becomes engaged during passive periods because such periods facilitate internal mental activity of the nature discussed above. Yet the network also activates during a variety of task-related contexts, particularly when task demands necessitate that individuals deliberately turn their attention inwards\(^32\). Along a similar vein, task contexts in which individuals experience internally-guided thoughts unrelated to the task at hand (e.g. planning what to cook for dinner) also engage the DN\(^33\). Of note, many regions throughout the DN also become engaged when participants experience a variety of positive and negative emotions\(^34\), yet their precise roles in emotion specifically, versus the thought processes that may often accompany such emotions, remain unclear. Both emotion and thought are likely constructive experiences, influenced by a variety of contextual, mnemonic, and conceptual factors\(^35\) – all processes previously linked to the DN, as we discuss below.

Converging evidence from resting state functional connectivity and task-related functional Magnetic Resonance Imaging (fMRI) studies suggests that the DN can be parcellated into separate components that may support different properties of internally-guided thought. These
components include a ventrally-positioned medial temporal lobe subsystem (DN_{MTL}), a more dorsally-positioned dorsal medial prefrontal cortex (DN_{DMPFC}) subsystem, and a centrally-positioned core (DN_{CORE})\textsuperscript{36} (Figure 1B). The MTL, and its cortical connections within the MTL-subsystem, have long been appreciated for their role in episodic memory\textsuperscript{37}. However, more recent evidence extends their mnemonic functions to other types of mental simulation such as episodic future thinking (also called prospection), and imagination more broadly\textsuperscript{38}. According to the constructive episodic simulation hypothesis, the DN_{MTL} allows us to bind together bits and pieces of past experiences and conceptual knowledge in flexible and adaptive ways\textsuperscript{39}. These operations are proposed to account for the re-constructive nature of our memories, and also allow for mental simulation of novel experiences – including hypothetical future ones\textsuperscript{38}. Activity within the DN_{MTL} often tracks the amount of visuospatial detail associated with mental simulation\textsuperscript{36,40}, suggesting it may play an important role in the specificity and visuospatial properties of memory, prospection, and other types of thought. The DN_{MTL} appears tied to more specific and contextually-detailed ways of thinking, which according to the construal level framework of Trope and Lieberman, would be referred to as a lower level of construal\textsuperscript{41}.

In contrast to the imaginative functions of the DN_{MTL}, the dorsal medial subsystem becomes engaged during thoughts characterized by higher levels of construal – that is, thoughts that are less detailed and more abstract in nature\textsuperscript{14,42}. The DN_{DMPFC} has been most strongly linked to tasks requiring individuals to reflect on their thoughts, memories, and emotions, as well as the mental states of other people – a process known as mentalizing\textsuperscript{14,43}. The DN_{DMPFC} is often recruited along with the DN_{MTL} during autobiographical memory and autobiographical prospection tasks\textsuperscript{44}, but it’s activity during such tasks has been linked to metacognitive reflection on the contents of the memory or mental simulation\textsuperscript{45}. Further supporting its role in abstract styles of thinking, including those involved in mentalizing, the DN_{DMPFC} is closely linked to brain systems supporting language, comprehension, and narrative processing\textsuperscript{14,46}.

Both subsystems are strongly interconnected with the DN_{CORE}, a set of mainly midline regions activating across a broad range of internally-guided tasks, especially when they involve referencing information to one's self\textsuperscript{14,47}. Activity in this network often signals strong perceptions of self-relevancy, personal affective salience, or interpersonal closeness on the
dimension of psychological distance\textsuperscript{48,49}. As such, the DN\textsubscript{CORE} may be especially engaged when individuals experience self-focused thinking, as expanded in Section 3. However, as with the DN\textsubscript{DMPFC}, the precise functions of the DN\textsubscript{CORE} are unclear and often debated. In fact, a recent investigation of network properties using high resolution neuroimaging revealed a more complex and topographically-interwoven architecture among DN components, calling the existence of a separate DN\textsubscript{CORE} into question\textsuperscript{50}. Additionally, it’s important to note that most of the studies referenced above have adopted analysis approaches assuming stability of neurocognitive processes over time. Recent inquiries into the dynamics of neural activity and connectivity reveal that DN components vary in their coherence with each other in flexible and contextually-dependent ways\textsuperscript{47}, see also \textsuperscript{48} – an observation that will be important to keep in mind when considering the dynamics of thought (discussed below).

3. INTERNAL EXPERIENCE GONE AWRY

The picture that emerges from the previous section is a promising one, suggesting that internal experience may bring about positive constructive functions that help us make sense of who we are now, where we want to be in the future, and what steps we need to take to get there. Activity in the default network may support many functional aspects of internal experience, spanning mental simulation to self-reflection. With so much promise, then, why might our capacity for introspection sometimes feel like a double-edged sword?

While many individuals often experience benefits and costs to our internal thoughts, this cost-benefit ratio varies widely between individuals. On the negative extreme are people who experience a constant internal struggle with their thoughts and emotions. In these cases, internal experience can take a significant toll on their personal mental and physical health, often interfering with the quality of their interpersonal relationships and occupational success. Many others fall somewhere in the middle of the spectrum from functional to dysfunctional.
Understanding the neurocognitive mechanisms driving this inter-individual variability might provide keys to unlocking introspection's full potential for beneficial change.

Dysfunctional thoughts are a hallmark of numerous mental health disorders. For example, rumination is characteristic of major depressive disorder (MDD), worry is characteristic of generalized anxiety disorder, and intrusive thoughts also plague those with obsessive-compulsive disorder (OCD) and schizophrenia. Prior studies suggest that there are many ways in which such thoughts can go awry. These include alterations in aspects of thought content (including affective content), as well as the processes by which individuals become aware of their thoughts, relate to their thoughts, and regulate how their thoughts arise and unfold across time. According to the content regulation hypothesis, the content characterizing off-task thoughts is an important factor constraining the cost and benefits of off-task thinking. This hypothesis proposes that individuals with poor well-being will show difficulty maximizing positive, constructive off-task thoughts, and minimizing negative, unconstructive off-task ones. Different disorders can be characterized to some degree by different content profiles, in line with the cognitive content specificity hypothesis of Beck and colleagues. There is considerable overlap in content across disorders as well, indicative of transdiagnostic processes at play. Depression and anxiety are considered characterized by overly negative thoughts, whereas mania appears to represent a focus on overly positive and grandiose content. Thoughts of loss and failure, and heightened past-focus seem to be a core feature of depression, as well as heightened self-focus in thought and language, which may overlap to some extent with symptoms of anxiety. Individuals with OCD and specific phobias exhibit thoughts with heightened negative content associated with a fear of personal harm. Of note, most existing studies investigate overall thinking patterns without distinguishing between instances in which thoughts are on-task versus off-task, marking an important avenue for future research. A few studies have characterized off-task thoughts in depression, highlighting more negative, self-focused, and past-oriented content. Additionally, most studies tend to either focus on a specific aspect of content (e.g. valence or temporal-orientation), examine the nature of participants' thoughts in laboratory contexts that may poorly mimic the complexities of daily life, or capture thinking patterns with a single trait questionnaire rather than
on numerous occasions in the moment. We propose that the clinical field could benefit from extending this inquiry to real world thought patterns across a variety of on-task versus off-task contexts using ecological momentary assessment techniques. Quantifying and comparing the full phenomenological profile of daily thinking patterns may help clinicians identify transdiagnostic mechanisms that could provide clues to effective therapeutic change. To the degree that these daily thinking patterns change prior to reported improvements in mood symptoms, one may be able to deduce that the patterns reflect causal mechanisms, and not simply correlates, of change.

In addition to alterations in thought content, a variety of mental health disorders exhibit alterations in processes that govern how individuals relate to, or connect with, their thoughts. Although thinking about personally-significant topics may help us find solutions to ongoing personal concerns, many individuals with poor well-being tend to believe that their momentary thoughts and emotions define them as individuals, a phenomenon referred to as "cognitive fusion". These individuals have difficulty psychologically distancing themselves from their thoughts, which loom large and dominate their attentional focus, and likely have stronger linkages with their affective states (Andrews-Hanna et al., unpublished data). People who score high on measures of cognitive fusion often take their thoughts and beliefs – especially about themselves – to be literally true (e.g., I am un-loveable, or I can’t figure this out). As a result, they tend to act on such beliefs by engaging in avoidance, compulsive, or other dysfunctional behaviors. Conversely, individuals scoring high on trait mindfulness show a heightened ability to separate themselves from their thoughts and emotions (e.g., I have moments when I don’t feel loved, or I can’t figure this out right now). By perceiving thoughts as transient experiences that come and go, mindful individuals are better able to adopt an accepting and non-judgmental stance towards their thoughts, and refrain from considering the contents of their thoughts as true reflections of their self-identity.

In summary thus far, a number of different mental health disorders appear to be characterized by alterations in the content characterizing daily and off-task thoughts, as well as a variety of processes influencing the degree to which individuals relate to their thoughts. Below we discuss a transdiagnostic style of thinking called maladaptive repetitive thought (RT) which, in addition
to the alterations described above, is characterized by another important dysfunctional thought process: regulations in the way thoughts arise and unfold dynamically over time.

3a. A Transdiagnostic Phenomenon - Maladaptive Repetitive Thought

A dysfunctional style of thinking called *maladaptive RT*, or sometimes *perseverative cognition*, has gained recognition as a transdiagnostic phenomenon\(^{69-72}\). Central to maladaptive RT is its recurrent or perseverative nature, signaling restricted dynamics that may manifest as reduced variability in thought content over short or extended durations. Although this variability in thinking is not often quantified experimentally, there is some suggestion that individuals who score high on questionnaires assessing aspects of maladaptive RT think about fewer topics compared to healthy controls, particularly when experimentally primed to be in a state of ruminative self-focus\(^{73}\).

Although not all repetitive thoughts may be dysfunctional, RTs experienced as intrusive or difficult to control, and predominantly negative in content, appear to have maladaptive consequences on health and well-being\(^{12}\). For example, such RTs contribute to the onset and maintenance of mental health disorders over time\(^{74}\), and are associated with heightened physiological reactivity, poor sleep quality, and a range of other cardiovascular, autonomic and endocrine disruptions\(^{75,76}\). Maladaptive RTs prolong stressful past, present, or future experiences by maintaining cognitive representations of those experiences over time, leading to sustained physiological states of preparation or readiness\(^{75}\). As such, maladaptive RT is thought to be a mechanism by which acute stressors can have chronic physiological effects, increasing the risk for a number of diseases\(^{75,77}\).

Maladaptive RT may signify impaired or over-taxed attentional control processes that might disrupt attention to the task at hand and manifest as elevated levels of off-task thinking. Indeed, ruminative individuals are more likely to engage in off-task thinking in laboratory settings and in daily life\(^{62,78}\), and off-task thinking is more likely when negative, self-relevant information is made more salient in mood or rumination induction paradigms\(^{79}\). Alterations in the context of off-task thinking in individuals who exhibit maladaptive RT are in line with *context regulation*
hypotheses, which propose that impairments in the ability to restrict off-task internal thoughts to easy or unimportant task contexts, may lead to maladaptive functional outcomes on behavior and well-being\textsuperscript{11,12}. For example, the rumination as avoidance hypothesis states that repetitively focusing on the causes of a loved one’s death can prevent awareness of the painful feelings resulting from the permanent loss of the person\textsuperscript{80}.

\textit{3a1. Subtypes of Maladaptive RT}

Within the broader construct, there are several specific types of maladaptive RT commonly discussed in the clinical science literature. \textit{Rumination} is associated with current depressive symptoms\textsuperscript{81,82} and prospectively predicts depression onset, intensity, and duration\textsuperscript{67,83}. Yet rumination is also evident in other disorders, including most prominently, anxiety\textsuperscript{67}. Rumination is characterized by negative and often self-focused content, and involves perseveratively fixating on one's symptoms and causes of distress\textsuperscript{67}. Rumination is tightly linked to negative mood, and has been shown to both precede and follow negative affective states in laboratory settings and daily life\textsuperscript{12,64,84}.

\textit{Worry} is regarded as a common human experience independent of psychopathology \textsuperscript{66,85}, but excessive worry is a characteristic feature of many mood and anxiety disorders\textsuperscript{58}. Like rumination, worry is negative in affect and often difficult to control\textsuperscript{85}. This subtype of repetitive thought is predominantly verbal in form and future-oriented, pertaining to uncertain future situations feared to have negative outcomes\textsuperscript{85,86}. Worry is often used as an unconstructive attempt at problem solving salient concerns, an avoidance strategy for confronting threatening stimuli, and as a way to mentally prepare for the "worst-case-scenario." However, over time, worry becomes negatively reinforced – exacerbating and prolonging negative affect and arousal, with severe long-term health consequences\textsuperscript{75}. Behaviorally, worry is also associated with procrastination, leading to difficulties in reaching one's goals\textsuperscript{87}.

\textit{Obsessions} are typically associated with OCD, and involve recurrent, intrusive, and disturbing thoughts about current experiences or events. Obsessive thoughts are usually followed by
compulsions, and frequently the person maintains that their thoughts or behaviors exert an exaggerated degree of control or causal influence over the world\textsuperscript{88}.

Recently, two additional subtypes of maladaptive RT have been proposed\textsuperscript{72}. \textit{Yearning}, or the unrelenting desire for someone\textsuperscript{89}, is a hallmark symptom of the newly minted \textit{Persistent Complex Bereavement Disorder}\textsuperscript{57} or \textit{Prolonged Grief Disorder}\textsuperscript{90}. These perseverative, intrusive thoughts can interfere with developing current life goals and meaning, leading to dysfunction and significant emotional distress. \textit{Interoceptive repetitive thought} is the uncontrollable and repeated attending to one’s bodily sensations (e.g., breathing or digestion) and maladaptive appraisal of these sensations\textsuperscript{72}. This somatic hypervigilance is prevalent in somatoform and anxiety disorders\textsuperscript{91,92}. Collectively, these observations suggest that interventions that help individuals to either engage in more positive constructive thought content, strengthen attentional control, improve awareness of thoughts, or alter the relationships people experience with their thoughts, may help improve well-being. We turn to such therapies in Section 6.

As illustrated in the transdiagnostic phenomenon of maladaptive RT, dysfunctional thought across numerous mental health disorders often manifests as alterations that extend beyond that of thought content alone. The repetitive nature of maladaptive RT suggests additional alterations in the processes that regulate how thoughts are initiated and unfold over time. Although the precise nature of the mechanisms that regulate the dynamics of internal thought remain to be fully understood, we recently put forth a neurocognitive model highlighting these processes\textsuperscript{6} (see also\textsuperscript{8,93}), and unpack it in the next section with the hope it may provide an illuminating clinical framework.

4. SPONTANEITY AND CONSTRAINTS ON THOUGHT: A DYNAMIC FRAMEWORK

According to the dynamic framework model, an important characteristic underlying the varieties of internal thought is the set of neurocognitive mechanisms that constrain the contents of one's mental state, and the transitions between one mental state to another. When constraints on thought are low, \textit{spontaneous thoughts} are likely to emerge. Spontaneous thoughts unfold and flow in a flexible manner, transitioning with ease from topic to topic. Conversely, when constraints are high, thoughts become more rigid over time, characterized by a narrower
conceptual scope with fewer dynamic transitions. Neurocognitive constraints may be mechanistic because when they are imposed, they may help stabilize a particular pattern (or mode) of distributed neural activation over time. The model highlights two classes of constraints on thought, each with different neural underpinnings and different clinical implications.

4a. Deliberate Constraints

Deliberate constraints are evoked and maintained deliberately, using top-down control mechanisms to allocate attention to specific sources of information, and to help buffer such attention from distraction. Deliberate constraints are exerted during goal-directed cognition, and can be directed towards external stimuli available to the senses, or to internal representations spanning thoughts, emotions, memories, plans and mental images (Figures 2 and 3A).

The majority of our knowledge about the neurocognitive mechanisms underlying deliberate constraints comes from contexts in which individuals deliberately direct their attention towards task-relevant information in the context of experimentally-constrained tasks94. However, deliberate constraints can also be evoked in an off-task manner, as when individuals intentionally choose to direct their attention away from the task or activity at hand95. In these contexts, what constitutes a "task" is tricky because sustained attention toward intrinsically-motivated information may reflect individuals' efforts to re-define and re-prioritize tasks, perhaps according to a hierarchy of intrinsic and extrinsic goals. Off-task attention can be directed externally (towards external, task-independent stimuli) and/or internally (towards task-independent thoughts and emotions). For example, when listening to a dull lecture, a student may deliberately decide to re-direct his/her attention toward the second-hand ticking around the clock. Likewise, while participating in a psychology experiment, a participant may deliberately disengage from the task to mentally prepare for her upcoming Physics exam. In both cases, the
way in which these mental states unfold over time should be relatively constrained, limited in scope to the topics under the deliberate focus of attention.

Most prior studies have lumped deliberate and non-deliberate (i.e. unintentional) off-task thoughts under the broader category of off-task thinking. However, recent research suggests this distinction may have important implications for mental health and cognitive functioning⁹⁵. Individuals high in trait mindfulness report more frequent deliberate off-task thinking⁹⁶, while individuals with symptoms of OCD and attention deficit hyperactivity disorder (ADHD) endorse a higher frequency of unintentional off-task thinking⁹⁷,⁹⁸. However, as we discuss below, unintentional thoughts can be further classified into spontaneous thoughts free from constraints, and thoughts that are highly constrained by automatic (often affective) processes. We hypothesize that automatically-constrained thoughts may largely account for the maladaptive associations with unintentional thinking.

Deliberate constraints are thought to be implemented by a set of brain regions forming the frontoparietal control network (FPCN)⁹⁹ -- one of seven large-scale networks in the 7-network functional parcellation of Yeo and colleagues¹⁰⁰ (Figure 1A and C). The FPCN includes at least two subsystems, often referred to as the cingulo-opercular network and the frontoparietal network, which may implement control on different timescales¹⁰¹. The FPCN as a whole, and the frontoparietal subsystem especially, seems to function as a brain-wide network hub, showing the highest degree of functional connectivity with numerous other large-scale brain systems at rest¹⁰². During experimentally-directed tasks, the FPCN is thought to dynamically reconfigure its coupling with other large-scale networks to allocate attentional resources to externally or internally-oriented goals⁵². For example, during visuospatial planning tasks, the FPCN increases its coupling with the dorsal attention network (a network supporting externally-oriented attention) and decreases its coupling with the default network. Yet when individuals plan their personal future goals, the FPCN reverses its pattern of coupling¹⁰³. Therefore, in the context of our model, we suspect that the nature of functional coupling with the FPCN should depend on the nature of the mental state the individual intends to promote or prohibit.

4b. Automatic Constraints and Affective Salience
In contrast to deliberate constraints, in which individuals intentionally guide their thoughts over time using executive control, individuals may find that their thoughts sometimes emerge and unfold in the absence of their deliberate control. The dynamic framework model differentiates between two different classes of unintentional (i.e. non-deliberate) thoughts. One class we discuss below is influenced by automatic constraints, which capture and draw one's attention to salient sources of information.

Like deliberate constraints, automatic constraints can apply to external stimuli or internal representations, but automatic constraints can act faster, sometimes occurring early enough to be outside of conscious awareness. *Sensory salience* mechanisms are at play when salient external stimuli capture and hold one's attention over time\(^6\). For example, someone may find herself continually distracted by a person talking loudly across the room, or surprised when the nature or location of upcoming external stimuli deviates from her expectations, signaling sensory prediction errors. Of relevance to numerous mental health disorders, salient internal representations – spanning thoughts about topics deemed to be of strong personal significance or motivation, intense emotional feelings, and/or heightened physiological responses – capture and constrain the nature of one's internal experience, restricting the scope and progression of thoughts and feelings over time\(^{104}\) (Figure 3B). Within this framework, states of maladaptive RT such as worry, rumination, or obsessive thoughts are characterized by excessive automatic constraints because these states are associated with strong biases to think and feel in a certain way, limiting the flexibility and dynamics of thought. In many situations, external and internal salience interact to guide behavior, as when threatening external stimuli evoke *affectively salient experiences*\(^{105}\), which drive motor or visceral-autonomic responses\(^{106}\).

An important aspect of our dynamic framework is that not all information will receive similar attentional priority across contexts within the same individual, or across different individuals with different patterns of thinking and feeling. An individual's current psychological state – including their mood, motivational goal state, or homeostatic drives -- can substantially influence what stimuli or internal representations will receive attentional priority, and thus be deemed as "salient"\(^{107,108}\). The influence of these contextual effects on attention may also differ across
individuals, such that a person who experiences more frequent or intense negative mood states (as in many mood and anxiety disorders) may experience negative material as more salient.

Consistent with the integrated memory model of Lane and colleagues\textsuperscript{109}, automatic constraints can be shaped by past experiences and previously-learned associations, which can, in turn, drive new learning\textsuperscript{110}. By this account, stimuli deemed salient in the past may be more likely to be salient now and in the future, given emotional biases in memory and attention that may become more pronounced if individuals ruminate on this information in the meantime. For individuals with maladaptive RT, maladaptive patterns of thinking become learned\textsuperscript{85}, leading to more potent affective salience systems, that over time may gradually take control of one's internal experience over a wider range of stimuli and contexts, making it easier to generalize to novel situations and lead to overgeneralization and catastrophic thinking patterns. As such, when a salient (i.e. negative or self-focused) thought is initiated, we hypothesize that it activates these learned negative associations, leading to trains of thought that last longer and are less flexible in their transitions from one thought to another. This cognitive inflexibility -- reflected as a lack of dynamic shifts in affective content -- may be particularly noticeable when an individual is already in a salient affective state, or when she is confronted with particularly salient information\textsuperscript{111,112}.

Automatic constraints are likely implemented by numerous dynamically-interacting brain networks that signal salient sensory, interoceptive, emotional, habitual, and/or personally-relevant information. Two networks that have been frequently-discussed in the literature are the \textit{ventral attention network} (VAN) and the somewhat-overlapping \textit{salience network} (SN). The VAN is predominantly right-lateralized, with key nodes in the right temporoparietal junction and the right anterior insula (aI), extending into inferior prefrontal gyrus\textsuperscript{113}. The VAN has been shown to play a role in automatically re-orienting attention to salient sensory stimuli, of which visual stimuli have been researched most extensively. The \textit{salience network} is a bilateral set of brain regions, mostly discussed in relation to the aI and the dorsal anterior cingulate cortex (dACC)\textsuperscript{114}. Like the VAN, the SN signals salient external stimuli, but also extends to salient affective, interoceptive, and other internal representations. In the 7-system parcellation of Yeo and colleagues, the VAN and the SN cluster together into a single brain system\textsuperscript{100} (Figure 1A
and C), consistent with domain-general accounts of an overarching functional system involved in processing multiple types of salient information\textsuperscript{110}. Through its dense interconnections with other cortical and subcortical structures including the posterior/mid-insula, amygdala, hypothalamus and periaqueductal gray, the SN is well-poised to rapidly detect afferent interoceptive signals arising from the vagus nerve, and send descending signals via downstream autonomic pathways to coordinate visceromotor responses\textsuperscript{110}.

The DN\textsubscript{CORE} may also implement automatic constraints, given its important role in self-related processing (Figure 1B and C). The DN\textsubscript{CORE} is strongly connected to the wider DN, and may function as one of the brain's broader "hubs"\textsuperscript{115}. As reviewed in Section I, the DN\textsubscript{CORE} becomes engaged across a range of internally-guided paradigms, spanning affective, mnemonic, prospective, conceptual, and social domains, especially when tasks involve intentionally or unintentionally referencing information to oneself\textsuperscript{14}. The DN\textsubscript{CORE} is also robustly engaged during periods of awake rest\textsuperscript{29,32}, when individuals often think about self-relevant topics. Activity in the anterior medial prefrontal cortex, extending into rostral ACC, increases when individuals make judgments about other people, particularly when those judgments pertain to other people to whom we feel close\textsuperscript{116}. These and other findings suggest that activity in the DN\textsubscript{CORE} may be indicative of a strong sense of personal and emotional connection to a particular stimulus or train of thought. Integrating these findings with our model, we suspect that the DN\textsubscript{CORE} may implement automatic constraints on thought by biasing attention to salient self-relevant stimuli or internal representations, which may function to restrict dynamics and stabilize thought over time.

\textbf{4c. Spontaneous Thought}

Thus far, we have introduced two neurocognitive mechanisms that constrain the contents and dynamics of one's internal experience. Although we have discussed goal-directed cognition and maladaptive RT as phenomenological experiences that manifest when deliberate and automatic constraints are respectively high, our model considers deliberate and automatic constraints along two separate axes that can be evoked in a gradient-like fashion. At the other end of the spectrum is a class of cognition known as \textit{spontaneous thought}. 
Spontaneous thoughts are characterized by relatively minimal deliberate and automatic constraints on both the content of thought and on the transitions from one thought to another. The model considers three types of spontaneous thought that vary in the degree of constraints imposed on them. Dreaming is a normal type of spontaneous thought characterized by the least amount of constraints, while creative cognition is a class of spontaneous thought characterized by the most constraints, involving dynamic shifts between spontaneity and deliberate forms of thinking. In between dreaming and creative cognition is a common type of spontaneous thought called mind-wandering. Since mind-wandering differs from goal-directed cognition in the degree of constraints imposed, mind-wandering should manifest psychologically in terms of having more variability over time (Figures 2 and 3A). It's important to note that this dynamic process definition of mind-wandering differs from most prior definitions that consider mind-wandering purely by its off-task contents. According to the dynamic framework, an individual who deliberately disengages from the task at hand to plan his weekend is not engaging in mind-wandering, but rather imposing deliberate constraints to restrict the scope and flow of thought. Indeed, recent studies suggest that the on versus off-task nature of one's thoughts and their freedom to move are independent constructs. Defining mind-wandering by its processes rather than by its off-task content also illuminates differences between mind-wandering and maladaptive RT. Whereas content accounts of mind-wandering would consider maladaptive RT as negatively-valenced mind-wandering, the dynamic framework holds that maladaptive RT is not mind-wandering because affective automatic constraints make it difficult for the mind to move freely, or wander, as is the case with spontaneous thought (see also). In fact, the use of “repetitive” in RT suggests its constrained nature.

Accumulating evidence suggests that the DNMTL may support the emergence of spontaneous thought, providing a major source of variability in content. As reviewed in Section 1, activity within the DNMTL is high during periods of awake rest, when demands on thought are typically low, and spontaneous memories and other internal thoughts are most likely to occur. During these periods, individuals with greater connectivity within the DNMTL report engaging in more frequent mental simulation of the past and future. Although it is possible that these findings may be driven by deliberate rather than spontaneous autobiographical thinking (as
the distinction was not assessed in the study), several other studies also hint at the role of the 
DNMTL in spontaneous thinking. For example, expert meditators with extensive training in 
metacognitive awareness demonstrate high activity in the DNMTL during instances corresponding 
to the emergence of spontaneous thought119. People are also less likely to be aware of 
spontaneous thoughts than deliberate thoughts95, and a recent fMRI study using experience 
sampling probes to assess on and off-task thought showed greater DNMTL activity during 
unaware off-task periods compared to aware off-task periods120. Dreaming -- characterized by 
very few constraints on thought in our model6 -- has been shown to engage the DNMTL to a 
greater degree than awake rest, paralleled by decreased activity within the FPCN likely reflecting 
relaxed deliberate constraints121. The DNMTL has also been shown to play an important role in 
creative cognition, whereby early, more flexible stages of the creative process are associated 
with increases in activity in the MTL, while later, more deliberate and evaluative processes are 
linked to activity throughout the FPCN and other DN subsystems122,123.

The dynamic framework predicts that spontaneous thoughts and associated patterns of network 
connectivity (especially within the DNMTL) will exhibit more variability over time than deliberate 
thoughts. Conversely, deliberate thoughts would be more likely than spontaneous thoughts to be 
supported by enhanced activity within the FPCN, and enhanced connectivity between the FPCN 
and the DN, including the MTL subsystem. Supporting these predictions, individuals reporting 
higher scores on a trait daydreaming questionnaire demonstrated enhanced variability over time 
in DNMTL connectivity during periods of awake rest. Additionally, Golchert and colleagues124 
found that individuals who reported more deliberate task-unrelated thought on a trait 
questionnaire exhibited greater functional coupling between the FPCN and the DN, and greater 
cortical thickness in aspects of the FPCN. Finally, in the context of memory retrieval, 
involuntary episodic memories have been linked specifically to the DNMTL, while voluntary 
episodic memories additionally involve aspects of the FPCN, likely supporting deliberate control 
imposed on the retrieval process125. In summary, although uncovering the neural correlates of 
spontaneous cognition in experimental contexts is challenged by the lack of meta-awareness 
sometimes associated with spontaneous thinking, a variety of studies using indirect approaches 
to estimate spontaneous processes point to the critical role of the DNMTL.
4d. Interactions Among Types of Thought

An important consideration of our model is that it allows for interactions between, and shifts in, the processes that influence how thoughts are evoked and unfold over time (Figure 3). At one point in time, an individual's mental state or train of thought may be characterized as spontaneous, perseverative, or goal-directed in nature, depending on the presence or lack of constraints outlined previously. Yet seconds later, her mental state may shift in our state space model. For example, an individual's chain of thinking in an epoch of time could be characterized as spontaneous and freely-moving with ease. She may then become aware of her thinking, and implement deliberate constraints to guide her train of thinking in one way or another. These deliberately-constrained thoughts may then shift to spontaneously-emerging thoughts after constraints start to wane over time. Or perhaps an anxious individual notices he is worrying about an upcoming social confrontation that he has been fearing, which triggers perseverative, negative self-evaluative thoughts and a pounding heartbeat. He may then try to implement deliberate constraints to reappraise the situation or redirect his attention to his surroundings.

These examples illustrate the tension that sometimes exists between the types of cognition outlined in our model. Often, deliberate constraints are implemented to deliberately redirect attention away from affectively-constrained or spontaneously-evoked thoughts. The use of deliberate constraints to attempt to regulate one's thoughts and emotions is of particular relevance to a number of psychotherapeutic interventions, as discussed in the next section. Being aware of the nature of one's mental state may be an important process influencing these processes.

5. CLINICAL IMPLICATIONS OF DYSFUNCTIONAL CONSTRAINTS

The model outlined above offers testable predictions regarding the clinical consequences of neurocognitive alterations in deliberate, automatic, and spontaneous internal experiences.
Although comparison across existing studies is complicated by methodological and participant variability, these predictions are broadly supported by existing literature, and point to key mechanisms of therapeutic change.

One important prediction of our model is that individuals who exhibit strong impairments in the structural and/or functional integrity of the FPCN should exhibit difficulty deliberately controlling the nature of their internal experience. These impairments may manifest as increased distractibility from task-irrelevant stimuli or unrelated thoughts, poor task performance, and deficits in learning and memory. Deliberate off-task thoughts may also be less frequent, shorter, and more fragmented, and goal-directed internally-guided processes such as autobiographical planning and problem solving may be particularly compromised. As deliberate constraints restrict the dynamics of thought, an impaired FPCN predicts more temporal variability in the phenomenological characteristics of internal experience.

In line with these predictions, numerous mental health disorders including MDD, anxiety disorders, OCD, schizophrenia, bipolar disorder, and ADHD have been associated with structural impairments and/or functional abnormalities within the FPCN (reviewed in 126). In MDD and rumination, these findings include reductions in regional grey matter volume within the FPCN, weaker functional connectivity between FPCN regions at rest127, and more variable connectivity between the FPCN and the DN_{CORE} over time128, consistent with the FPCN providing less stable deliberate constraints on the flow of thought. Task-related hypoactivity of the FPCN is also frequently noted in MDD129, yet hyperactivity is also sometimes observed. In individuals with MDD and trait rumination, neural abnormalities in the FPCN are paralleled by impairments in major domains of executive functioning130,131, and an increased propensity for task-unrelated thinking62. Lending further support are studies showing improvements in MDD symptoms following transcranial magnetic stimulation of the dorsolateral PFC132, a key node of the FPCN. Collectively, these findings predict that clinical therapies that help to restore deliberate constraints on cognition by strengthening attentional control over one's internal experience may work particularly well.
Anxiety has been linked to somewhat similar alterations within the FPCN\textsuperscript{133}, yet alterations in salience networks are more consistently observed in anxiety\textsuperscript{133}. In anxiety and depression, negative and threat-evoking stimuli capture attention to a greater degree, which may occupy more working memory resources, leading to fewer resources available for additional cognitive processing\textsuperscript{134}. Considering the heightened salience of negative and threat-related material, stronger deliberate constraints may be required to suppress such material, leading to an overtaxed and ineffective FPCN. Whether dysfunctional deliberate control processes are a cause or a consequence of mood and anxiety disorders remains an open question.

Similar to deliberate constraints, automatic constraints should also stabilize internal experience over time, yet these constraints lack volitional control and work by automatically drawing one's attention to specific sources of information. For depressed, anxious, and ruminative individuals who have a history of experiencing maladaptive associations with certain classes of material, thoughts and attention may be more strongly drawn to this material, and consequently characterized by a narrower conceptual scope with less variability over time\textsuperscript{104}. In these individuals, our model predicts that brain networks supporting automatic affective constraints should be engaged more robustly and perhaps more frequently in mood and anxiety disorders, particularly when exposed to contexts that have previously evoked heightened attention.

Supporting these predictions, individuals with high state and trait anxiety exhibit heightened resting state connectivity within the salience network, and heightened salience network connectivity with the amygdala, a subcortical target of the network\textsuperscript{114,135}. These regions also demonstrate enhanced activity in anxious individuals during tasks involving negative and threat-related stimuli\textsuperscript{136}, and they are particularly engaged during panic attacks\textsuperscript{137}. Structural alterations in salience networks are also apparent, whereby increased trait anxiety is associated with white matter profiles indicative of more rapid and efficient network communication between the anterior insula and the basolateral amygdala\textsuperscript{135}. Behaviorally, these neural changes manifest as attentional biases to negative and threat-related stimuli that occur especially quickly, even prior to conscious awareness\textsuperscript{138,139}. At the pharmacological level, anxiolytics such as propranolol and alcohol have been found to dampen the physiological response to typically threat-provoking stimuli and reduce functional connectivity within the salience network\textsuperscript{140,141}. 
Although individuals with depression and trait rumination also exhibit alterations in activity and connectivity of the salience network\textsuperscript{142}, network alterations in these individuals are particularly pronounced in the DN\textsubscript{CORE} and the dMPFC. As reviewed in the previous section, activity in the DN\textsubscript{CORE} signals strong personal significance, and likely implements automatic constraints on cognition when external stimuli and internal thoughts are deemed particularly self-relevant, or psychologically "close" to one's current sense of self. Consistent with the model’s predictions, depressed and ruminative individuals exhibit heightened activity and functional connectivity within the DN\textsubscript{CORE}, both at rest and following rumination induction\textsuperscript{128,142–144}. Recent investigations into functional network dynamics in MDD using dynamic connectivity fMRI are consistent with these findings, showing stronger and more stable connectivity between the aMPFC and the parahippocampal cortex (a component of the DN\textsubscript{MTL}) at rest\textsuperscript{128}. Relatedly, enhanced dMPFC connectivity with other regions within and outside the DN has been observed in MDD\textsuperscript{142,145}, perhaps reflecting a predisposition towards more abstract, high levels of construal, including overgeneral (i.e. less specific) autobiographical memories\textsuperscript{146} and styles of thinking reviewed in \textsuperscript{53}, but see \textsuperscript{64}. Collectively, these findings predict particular promise for clinical therapies that dampen automatic constraints on cognition, increase the variability of internal experience, and cause the evaluation of core beliefs about the self using concrete thoughts and memories.

Exacerbated affective constraints on thought, and impairments in the ability to effectively implement deliberate constraints, should have downstream effects on the frequency of spontaneous thought and the integrity of its neural underpinnings, hypothesized to rely on an intact DN\textsubscript{MTL}. Although to our knowledge, spontaneous thoughts have yet to be investigated in MDD, anxiety, and other forms of maladaptive RT, we predict spontaneous thoughts will occur less frequently in these disorders, leading to detrimental outcomes on learning and memory, creativity, and mental health. Supporting these hypotheses, depression has been associated with reductions in hippocampal volume\textsuperscript{147}, accompanied by deficits in episodic and autobiographical memory\textsuperscript{146,148}. Compared to non-depressed individuals, depressed individuals have also been shown to demonstrate reduced DN\textsubscript{MTL} activity and enhanced anterior DN\textsubscript{CORE} activity during autobiographical memory and prospection tasks\textsuperscript{149}. 
6. MECHANISMS OF CHANGE

Across different psychotherapeutic modalities, maladaptive RT and other types of dysfunctional internal experience are addressed in a variety of ways. Sometimes the application of treatment interventions is very specifically designed to deal with the intrusive and distressing nature of maladaptive thoughts. At other times, the cognitive effect of psychotherapy is a byproduct of a therapeutic process targeting behavior or mood. These psychotherapeutic methods could be divided into three overarching strategies, although many other strategies might be distilled\textsuperscript{72,109}. They include: a) distracting individuals from their thoughts by promoting positive associations with behavior, b) manipulating the content of the thoughts themselves, c) promoting awareness and acceptance of thoughts. Below we introduce three classes of therapies adopting these strategies, and consider each’s potential for inducing enduring change from the perspective of the dynamic neurocognitive model outlined above.

6a. Behavioral Activation

An empirically-based psychotherapy called \textit{behavioral activation} (BA) is based on a model that highlights the association between mood, activity, and reinforcement\textsuperscript{150,151}. In depression, avoidance and withdrawal behaviors maintain depressive symptoms by preventing individuals from engaging in aspects of daily life that could have a rewarding, anti-depressant impact\textsuperscript{152,153}. BA strategies may include scheduling daily activities and rating the degree of enjoyment and satisfaction one derives from specific activities, among others\textsuperscript{152}. Strengthening positive associations with constructive behaviors may have lasting effects on clients’ perceived value of such behaviors, making it easier for them to prioritize such behaviors as deliberate “goals”.

Through repeated exposure to goal-oriented behaviors and increased feelings of self-efficacy, behavioral activation has been shown to successfully break down avoidant behaviors and maladaptive RT (including worry)\textsuperscript{154}. Although the objective in behavioral activation is to have clients engage in rewarding activities in their lives, an important secondary effect may be to distract them, or release them from the automatic constraints of maladaptive RT. This effect might manifest as an increase in cognitive flexibility, heightened variability in the content of everyday thought, and a more regular instantiation of goal-directed cognition.
6b. Cognitive Behavioral Therapy

Cognitive Behavioral Therapy (CBT) -- another empirically-based therapy with important links to our model -- is one of the most widely-practiced and extensively-researched clinical therapies, with evidence of long-term post-treatment benefits. Although CBT often incorporates some degree of behavioral practices, CBT is chiefly designed to improve well-being by targeting the maladaptive nature of dysfunctional thoughts and beliefs. CBT teaches clients to consider the accuracy and utility of their thoughts by examining the cognitive errors and flawed logic contained in such thoughts. Clients gain practice monitoring and labeling the content of their thoughts, the precipitating situations and dysfunctional beliefs that often evoke them, and the behavioral responses and mood changes that typically follow. For example, a person who developed a core belief through harsh parenting that they were too “big for their britches” might follow every thought of their own achievement with a self-deprecating internal comment. The person would likely be unaware of this sequence, due to the automaticity of the thought pattern. However, with repeated practice, negative semantic scripts and schemas that precipitate dysfunctional thoughts become modified, and automatic constraints on thought may gradually break down. Within the context of our model, everyday thoughts may become more positive, constructive, and less bothersome over time, leading to an increased likelihood of unfolding more easily and flexibly. The process of modifying thoughts and implicit conceptual associations may have downstream effects on one’s affective states and traits, improving mood and well-being over time.

Notably, initial neuroimaging studies examining the neural changes associated with CBT provide promising support for our model. CBT has been shown to reduce the neural reactivity and functional-anatomic integrity of salience networks and the DN CORE, while concurrently strengthening the structural integrity of the FPCN and hippocampus, and connectivity between the FPCN and DN. These results point to neural mechanisms underlying CBT that may allow individuals to evoke deliberate constraints to more effectively down-regulate their dysfunctional negative thoughts, and/or more rapidly and efficiently re-direct and sustain attention to more constructive forms of internal experience.
6c. Mindfulness-Based Therapies

Mindfulness as a process of regulating internal experience has been discussed throughout the present paper, and is a component of interventions such as Acceptance and Commitment Therapy (ACT)\textsuperscript{163}, Mindfulness-Based Stress Reduction\textsuperscript{164} and Mindfulness-Based Cognitive Therapy\textsuperscript{165}. In contrast to CBT, mindfulness approaches do not aim to modify the content of one’s maladaptive thoughts. Instead, these approaches work by helping individuals alter the relationships they experience with their thoughts, particularly by dampening the propensity for individuals to be overly “fused” with their thoughts, as discussed in Section 3. Mindfulness therapies teach individuals to let their thoughts emerge spontaneously without conceptual self-judgment, and to recognize that their thoughts are temporary, dynamic experiences that come and go and do not represent reflections of their self-identity – processes known as “defusion” or “decentering”\textsuperscript{65,68}. Support for these hypothesized mechanisms comes from studies demonstrating decreased activity within the DNe\textsubscript{CORE} in expert meditators\textsuperscript{166} and following mindfulness training\textsuperscript{167}. Given the possible role of the DNe\textsubscript{CORE} in implementing automatic constraints on thought by biasing participants towards personally-salient information, these findings suggest that the process of psychological distancing inherent to defusion may also serve to dampen personal automatic constraints and increase thought flexibility and variability. Future studies could test this hypothesis using dynamic neuroimaging and introspective techniques.

A critical step towards helping individuals change the relationships they experience with their thoughts is the process of becoming aware of those thoughts in the first place. Indeed, many exercises implemented in mindfulness therapies (i.e. meditation, yoga, breathing exercises, body scans, etc.) teach individuals to be more in tune with the present moment, spanning thoughts, emotions, perceptions, and sensations\textsuperscript{168}. When clients notice their attention has deviated away from the present moment, they are instructed to nonjudgmentally redirect their attention back to the here-and-now. These practices may function to strengthen attentional control and cognitive flexibility, providing individuals the means to become aware without judgment, and to flexibly shift their attention back to relevant “tasks” shortly after noticing the emergence of disrupting internal experiences.
The impact of mindfulness-based therapies on increasing present moment awareness, enhancing cognitive flexibility, and facilitating defusion may also relate to their promise for altering the level of construal with which internally-guided thoughts occur. Mindfulness therapies have been reliably shown to enhance the specificity of autobiographical memories\textsuperscript{169,170} and personal goals\textsuperscript{171}, suggesting an overall shift towards thinking characterized by lower levels of construal.

Within the model’s neurocognitive framework, we hypothesize that mindfulness therapies will rebalance the relationship between brain networks associated with automatic, deliberate, and spontaneous thought. For one, greater awareness of the present moment may counterintuitively increase activity within the anterior / mid-insula and the broader salience network. However, the process of decreasing absorption through defusion and non-judgment may shorten the temporal duration of salience network activity, and decrease the abnormally elevated activity and connectivity within the $D_{\text{CORE}}$ and between the $D_{\text{CORE}}$ and the salience network. Dampening the neural underpinnings that automatically constrain internal experience may create more opportunities for the emergence of purely spontaneous thought. These changes may also be paralleled by increases in the integrity and efficiency of the FPCN. Although mindfulness approaches do not aim to alter thought content, their promise for regulating the dynamics of thought may have unintended long-term effects on thought content, leading to an increase in the ratio of positive and constructive to negative, unconstructive thoughts over time.

Although initial studies provide some support for predictions of our model\textsuperscript{167}, comparison across studies is limited by the large degree of methodological limitations and heterogeneity. Additional support comes from meta-analyses of studies comparing experienced meditators versus novices, which may differ in a number of respects beyond meditation practice alone. Nonetheless, these studies highlight enhanced fMRI activity and changes in brain structure within the FPCN, insula, hippocampus, and other regions, and decreases in activity within the $D_{\text{CORE}}$\textsuperscript{167,172,173}. As with the other clinical therapies discussed above, it remains for future research to determine the robustness and specific mechanisms behind these findings.

7. SUMMARY AND FUTURE DIRECTIONS
Throughout this chapter, we have reviewed the intersection of basic science and clinical research on internal experience, highlighting a traditionally underappreciated facet of human cognition with important mental health relevance. In the first half of the chapter, we discussed evidence suggesting that in the absence of mental health concerns, task-unrelated and related thoughts show largely positive, constructive characteristics indicative of adaptive problem solving, planning, and social navigation. Neuroimaging studies suggest that these facets of self-generated thought are supported by a heterogeneous brain network called the default network. When mental health concerns are present, daily thinking patterns are more likely to be characterized as negative, unconstructive, and (often) overgeneral. Yet alterations extend far beyond changes in thought content alone. Dysfunctional thought patterns are often repetitive and uncontrollable, suggesting additional changes in ways thoughts arise and unfold over time. Maladaptive outcomes may be associated with alterations in the relationships people experience with their thoughts, with heightened connection or thought fusion, perhaps signifying stronger emotional and personal salience. Impairments in meta-cognitive awareness also often contribute to poor mental health, likely contributing to additional difficulty regulating one’s internal experience.

In the second half of the chapter, we introduced a recently developed dynamic neurocognitive framework for understanding the mechanisms of dysfunctional internal experience. This model highlights how automatic constraints on thought draw one’s attention to salient sources of information, restricting the conceptual scope and dynamics of thought over time. Deliberate constraints, which serve to stabilize internal experience using top-down control, can be evoked to overcome automatic constraints and guide thoughts in a volitional fashion. While clinical research is consistent with alterations in both automatic and deliberate constraints -- as well as the salience, DN\textsubscript{CORE}, and frontoparietal control networks that support these constraints -- little is known about the characteristics of spontaneous thought in healthy or clinical populations. Existing research suggests the DN\textsubscript{MTL} may play a critical role in eliciting spontaneous thought, which appears structurally and functionally altered in depression. The dynamic framework also makes several predictions regarding the mechanisms by which beneficial change might occur. Thus in the last part of the chapter, we survey common clinical therapies targeting different mechanisms, and note preliminary effects of therapy on brain activity and connectivity.
7a. Future Directions

The last several years have brought promising advances in our understanding of the brain and behavioral underpinnings of clinical dysfunction. Yet many questions still remain, and relatively little is still understood about how clinical therapies actually work\textsuperscript{174}. This chapter joins others in this edited volume to exchange cross-disciplinary perspectives on enduring change by bringing together basic scientists and clinical psychotherapists.

Here we propose that greater insight into the mechanisms of change should start with a deeper understanding of the nature of internal experience itself, including the content characterizing functional and dysfunctional thoughts across a variety of contexts, as well as the processes constraining how thoughts arise and unfold over time. To address these gaps, future research could include ecological momentary assessment paradigms to explore how multiple facets of thinking spanning the content and processes discussed above arise and unfold across a variety of real-world contexts. Neuroimaging studies could relate the content and dynamics of thought to the three brain systems hypothesized from our dynamic framework model (the DN, salience network, and FPCN) by examining how thoughts and their corresponding neural underpinnings arise, unfold and interact over time. Studies could examine the dynamic trajectories of thoughts as they occur naturally during unconstrained states, as well as before and after experimentally manipulating automatic and deliberate constraints on thought.

The predictions extending from our dynamic framework model (see Sections 4 and 5 and Figure 3) could be tested in individuals spanning the spectrum of functional to dysfunctional thoughts, both before and after treatment with clinical therapies. These efforts could be paralleled by developing new paradigms that seek to supplement introspective report and uncover less disruptive and multi-method indices of internal experience, particularly in the context of clinical interventions. Finally, insight into mechanisms of change could be advanced by measuring co-occurring physiological signals during single therapy sessions, by analyzing aspects of the dialogue between clients and therapists during such sessions, and by adopting exercises employed during therapy into specific tasks that can be mimicked in laboratory experiments. We hope that bringing together basic scientists and clinical psychotherapists to share their
perspectives on enduring change will inspire future collaborative research on this important topic.
GLOSSARY

**Automatically constrained thought:** Inflexible thoughts that occur outside of one’s deliberate control and that are biased towards a restricted set of information. Includes affectively salient thoughts, personally salient thoughts, habitual thoughts, maladaptive repetitive thought, and thoughts coupled to perceptually salient stimuli (e.g. a fly buzzing around the room).

**Default network:** A brain system comprised of cortical (mostly association regions), subcortical, and cerebellar regions that supports the emergence of internally-guided thought. Comprised of disparate components that contribute to different aspects of internally-guided thought. One DN component may bias thoughts towards personally-salient information, providing a source of automatic constraints; another DN component may support the emergence of spontaneous thought.

**Deliberately constrained thought:** Internally-guided and/or perceptually-coupled thought characterized as being evoked and maintained deliberately, using top-down control.

**Frontoparietal control network:** A brain system comprised of cortical (mostly association regions), subcortical, and cerebellar regions that supports deliberate, goal-directed cognition and provides a source of deliberate constraints on thought.

**Internally-guided thought:** Cognition that unfolds in a manner not directly tied to immediate perceptual stimuli. Considered synonymous with *internally-oriented thought, internal mentation, self-generated thought, perceptually-decoupled thought, stimulus-independent thought, or imaginative thought* (broadly construed).

**Maladaptive repetitive thought:** Automatically-constrained thought characterized by being restricted in focus, perseverative in nature, and associated with maladaptive consequences on well-being.

**Mental state:** As in Christoff et al.⁶: “A transient cognitive or emotional state of the organism that can be described in terms of its contents (what the state is ‘about’) and the relation that the
subject bears to the contents (for example, perceiving, believing, fearing, imagining or remembering).”

**Perceptually-guided thought:** Cognition that unfolds in a manner directly tied to information currently available to the senses. Synonymous with *perceptually-coupled thought* and *stimulus-dependent thought*.

**Resting state functional connectivity:** A neuroimaging technique examining temporal correlations in brain activity between spatially-disparate regions during extended periods of awake rest.

**Spontaneous thought:** Thought that arises and transitions relatively freely due to an absence of strong deliberate or automatic constraints on cognition.

**Salience Network:** A system of cortical, subcortical and cerebellar regions that implements automatic constraints on cognition, biasing thoughts towards perceptually and affectively-salient sources of information.

**Off-task thought:** Internally-guided and/or perceptually-guided thought whose content is unrelated to the task at hand. Synonymous with *task-unrelated thought*.

**Thought:** We operationalize “thought” as in Christoff et al.⁶ as “A mental state, or a sequence of mental states, including the transitions that lead to each state.” Our definition includes consciously-accessible thoughts, as well as thoughts that are not immediately accessible to conscious awareness.
ACKNOWLEDGEMENTS

The authors would like to acknowledge Zachary Irving for his extensive contributions developing the neurocognitive model highlighted in this chapter, as well as Richard Lane, Lynn Nadel, an anonymous reviewer, Kieran Fox, Nathan Spreng, and Matthew Dixon for additional helpful discussion. J.A.H. and M.-F.O. extend their additional gratitude to the University of Arizona Department of Psychology, the Cognitive Science Program, and the Arizona Alzheimer Consortium for supporting J.A.H and M-F.O. while preparing this chapter.
FIGURE CAPTIONS

**Figure 1.** *Key Brain Networks Contributing to the Dynamics of Internal Thought.*

A) As discussed in the text, several large-scale, interacting brain systems make important contributions to the dynamics of internal experience. The frontoparietal control network (FPCN; orange) is thought to implement deliberate constraints, providing a source of stability on thought. A broader salience network (SN; purple) is thought to implement automatic constraints, also contributing to thought stability. B) The default network (DN; pink) can be fractionated into subsystems that may contribute differently to the dynamics of internal experience. The DNCORE is hypothesized to provide automatic constraints on thought, while the DNMTL may allow spontaneous thoughts to emerge, providing a source of variability. C) Hypothesized network interactions associated with the dynamics of thought. Arrows represent functional connections between brain networks, shaded by whether they provide a source of variability (as in spontaneous thought), or constraints on thought. FPCN = frontoparietal control network; DNCORE = Core subsystem of the default network; DNMTL = medial temporal lobe subsystem of the default network; DAN = dorsal attention network, thought to support visuospatial attention. Figures adapted from Yeo et al.¹⁰⁰ and Christoff et al.⁶.

**Figure 2.** *A Neurocognitive Model of the Dynamics of Internal Thought.*

As described in the text, and as originally outlined by Christoff et al.⁶, the way in which internal thinking emerges and unfolds over time depends on the degree of deliberate or automatic constraints concurrently imposed. Spontaneous thoughts -- spanning dreaming, mind-wandering, and creative cognition -- are relatively free from constraints on thought. As a result, spontaneous thoughts are characterized by a high degree of variability in content over time. Deliberate and automatic constraints function to stabilize internal thought over time, restricting the conceptual scope and dynamic flow of thought. Goal-directed cognition is characterized by strong deliberate constraints. Maladaptive repetitive thought is characterized by strong automatic constraints. Note that we conceive of deliberate and automatic constraints on a continuum, with smoother boundaries between classes of thinking than illustrated here. Figure reproduced with permission from Andrews-Hanna et al., in press⁸.
**Figure 3.** Hypothesized Neurocognitive Underpinnings of Transitions between Classes of Internal Thought. In this hypothetical example of an extended stream of thought, an individual may find that his/her thoughts transition between spontaneous, automatic, and deliberately-constrained internal thinking. Functional network interactions hypothesized to support each class of internal experience are illustrated below this first-person example of subjective experience. Thicker arrows represent stronger functional connections between networks. FPCN = frontoparietal control network; DN\textsubscript{CORE} = Core subsystem of the default network; DN\textsubscript{MTL} = medial temporal lobe subsystem of the default network. Figure adapted from Christoff et al.\textsuperscript{6}. 


REFERENCES


10. McMillan RL, Kaufman SB, Singer JL. Ode to positive constructive daydreaming. *Front


55. Beck R, Perkins TS. Cognitive Content-Specificity for Anxiety and Depression: A Meta-


89. Robinaugh DJ, Mauro C, Bui E, et al. Yearning and Its Measurement in Complicated


156. Beck JS. *Cognitive Behavior Therapy, Basics and Beyond*. 2nd ed. New York City: The


167. Farb NAS, Anderson AK, Segal ZV. The mindful brain and emotion regulation in mood


Sources of variability

Deliberate constraints

Automatic constraints

FPCN

SN

DN

DN-DMPFC

DN-CORE

DN-MTL
Automatic Constraints

Strong

Weak

Deliberate Constraints

Weak

Strong

Spontaneous Thought

Maladaptive Repetitive Thought

Night Dreaming

Mind-wandering

Creative Thought

Goal-Directed Thought

Highly Dynamic Stream of Thought

Content

Time

Highly Rigid Thoughts / Little Dynamic “Flow”

Content

Time
A Deliberately constrained thought

“While I step up the curb, I realize that my thoughts are making me miserable. I decide to think about something else. Where am I heading to? Oh yes, groceries! I imagine myself walking down each aisle. I should get eggs and milk from the refrigerated section, potatoes and cauliflower from produce…”

B Automatically constrained thought

“As I cross the street, I begin to worry about the story that my newspaper wants me to write before I leave to Tahoe. Can I submit on time? Will anymore read a piece on trade unions? I picture my scowling editor. Do I even belong here?”

C Spontaneous thought

“While I walk to the grocery store, I daydream about the winter boots I’ve ordered from an online store, recall that blustery winter when they shut down my elementary school, then envision next weekend’s ski trip to Lake Tahoe.”

--- Sources of variability ➔ Deliberate constraints ➔ Automatic constraints