From “Intrusive Thinking: From Molecules to Free Will,” edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

17

Interventions and Implications

Judson Brewer, Harriet de Wit, Aurelio Cortese,
Damiaan Denys, Colleen A. Hanlon, Emily A. Holmes,
Martin P. Paulus, Jens Schwarzbach, and Peter Tse

Abstract

This chapter provides a framework for developing interventions that specifically target intrusive events. It describes the challenges in defining intrusive thoughts and the difficulty in distinguishing normal processes of cognition and emotion from indicators of dysfunction, defined from practical, neurobiological, or cultural points of view. Throughout, the term *intrusive events* is used to encompass both thoughts and images that become intrusive. Examples are explored as they occur in different psychiatric disorders to demonstrate their variance in form, frequency, and controllability. Treatment modalities that have been used to alleviate intrusive events in different psychiatric disorders are reviewed, including behavioral, pharmacological, and emerging electromagnetic brain interventions. Two clinical vignettes illustrate the nature and severity of intrusive events in patient populations as well as the complex, multidimensional nature of the clinical reality. Ways of measuring intrusive events are examined and deconstructed into components (e.g., sensory, motor, and cognitive features). By examining intrusive events across diagnostic categories, common basic biobehavioral processes may be revealed which, in turn, could facilitate the study of neural processes underlying the behaviors. A model of cognitive and emotional decision making is presented to provide a basis for understanding and studying intrusive events. Examples of how the model might account for the "failure modes" in intrusive events are used to formulate testable hypotheses, and future interventions that combine multiple treatment modalities are considered. The chapter concludes with a discussion of the broader cultural context of intrusive events.

Group photos (top left to bottom right) Judson Brewer, Harriet de Wit, Aurelio Cortese, Martin Paulus, Emily Holmes, Judson Brewer, Damiaan Denys, Colleen Hanlon, Peter Tse, Jens Schwarzbach, Martin Paulus, Colleen Hanlon, Aurelio Cortese, Harriet de Wit, Judson Brewer, Peter Tse, Damiaan Denys, Martin Paulus, Harriet de Wit, Jens Schwarzbach, Emily Holmes

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

Introduction

Defining the Phenomenon

What is an intrusion? To which entity does the concept of an intrusion correspond? Is it a single concept that refers to one and the same mental phenomenon, or are there multiple types that arise under different circumstances and across multiple mental states? Are intrusions different in animals than in humans, and do they vary across cultural contexts or even time?

The definition by Clark (2005:4) serves as a starting point for our analysis:

> …unwanted, clinically relevant intrusive thoughts, images, or impulses [are] any distinct, identifiable cognitive event that is unwanted, unintended, and recurrent. It interrupts the flow of thought, interferes in task performance, is associated with negative affect, and is difficult to control.

Accordingly, an *intrusion* can be understood both as a clinical symptom and as a regular mental phenomenon. In as many as 85% of healthy individuals, for instance, intrusions have been observed in the form of thoughts, images, or impulses (Rachman and de Silva 1978), experienced as an individual event that is undesirable, unintentional, and recurrent. In comparison, intrusions as a clinical phenomenon may be fundamentally different. Their severity may increase over time, despite or even because of the interventions undertaken by the patient or caregiver, gradually consuming increased amounts of a patient's time and energy. In addition, pathological intrusions rarely diminish on their own. Unravelling the principle of "reinforcement" that leads to a snowball effect poses a challenge for both clinical and neurobiological researchers.

As suggested by Clark (2005) and confirmed in our discussions, future empirical studies need to clarify the boundaries of what constitutes an intrusion for psychiatric clinical and psychological purposes. While a *minimum definition* runs the risk of being too restrictive, and may not describe all types of intrusive phenomena, it permits us to distinguish between normal and abnormal intrusions. This, in turn, is needed by diagnosticians to distinguish between sick and healthy mental events and aid in treatment decision making. As a minimum definition, we propose the following:

> Intrusive events are unwanted, clinically relevant, intrusive thoughts, images, or impulses that an individual may attempt to resist, but which are out of their control.

A *maximum definition*, by contrast, needs to incorporate all intrusive phenomena, as exhaustively as possible, into a general descriptive approach. In this way, an intrusion will be able to be described as a specific mental event or experience and, regardless of its nature, be studied by different scientific disciplines. We propose the following as a maximum definition:

> An intrusive event is any interruption in the flow of mental events by an external (e.g., a ringing telephone) or internal (e.g., a thought) stimulus.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

Many types of mental events can intrude into consciousness. Verbal and nonverbal thoughts, mental images, impulses, memories, emotions, desires, and dreams can all reset the contents of consciousness and be experienced as unwanted or intrusive. Do all such events have the potential to be clinically relevant? How does the minimum definition, which primarily encompasses psychiatric symptoms, relate to the maximum definition, which relates to everyday experienced phenomena? To what extent do they overlap or deviate from each other? To what extent are they qualitatively, or perhaps only quantitatively, different from each other? Importantly, if clinically relevant intrusive thinking takes different forms or domains (e.g., verbal vs. imagery), what are the implications for treatment? Different treatment approaches may be needed or optimized for different domains.

In our discussions, we juxtaposed these two definitions next to each other—one with a psychiatric clinical purpose, the other with a psychological fundamental purpose—but wish to emphasize that intermediate viewpoints are possible. To address our group's topic, however, we found these artificially contrasting definitions helpful. A discussion of the philosophical and social implications of defining the phenomenon is included later in the chapter.

Intrusive Events and Psychiatric Disorders

Across many common psychiatric disorders, intrusive events present as key symptoms (see Schlagenhauf et al., this volume). Recurrent unwanted thoughts and images occur in almost every psychiatric disorder and are explicitly described among the criteria that must be met for a formal diagnosis in several disorders. Notably, intrusive events may be problematic symptoms in themselves. As such, they are appropriate targets for interventions but have rarely been identified as a transdiagnostic clinical feature constituting a target for treatment (cf. Iyadurai et al. 2018). Given their importance across mental disorders, intrusive events may offer insight into the neural mechanisms involved in the pathophysiology of psychiatric conditions.

In our discussions we de-emphasized intrusive events that are manifest in perceptual or thought disorders (e.g., hallucinations or intrusive delusions associated with schizophrenia or psychotic depression) or tics (as in Tourette syndrome). These may fall into separate categories of events and may notably lack the "negative affect" component specified by Clark (2005). As a result, our focus here is on the specific symptoms of intrusive events, not on full psychiatric diagnoses.

Intrusive events vary in their prominence as defining symptoms for different diagnoses. Importantly, the pathophysiology of intrusive events may differ across diagnostic categories and may differ from other symptoms within a disorder. For some disorders, such as posttraumatic stress disorder (PTSD) and obsessive-compulsive disorder (OCD), intrusive events appear to be central

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

defining features and may causally drive other symptoms (see Holmes et al. as well as Hanlon, this volume). In others, intrusive events may not be the defining feature of a disorder but rather one of many criteria used to reach a diagnosis. For example, craving is a symptom of substance use disorder (SUD), but SUDs can and do appear without cravings. Similarly, suicidal ideation appears in major depressive disorder, but the disorder can occur without it. By focusing on a single symptom, it may be possible to relate the clinical manifestation of an intrusive event to a dysfunction of neural circuits controlling normal brain function.

Given the heterogeneous nature of intrusive events associated with different psychiatric disorders, we recommend that a research program be developed to examine whether subcategories of intrusive events exist; this information is needed to establish an association with the underlying neurobiology. More importantly, intrusive events with a different neurobiological signature may require distinct treatment approaches. Given the different dimensions of an intrusive event (e.g., prior experience and expectancies, precipitating events, temporal sequence of events, emotional breadth, contextual features, psychological consequences), we offer two clinical vignettes to illustrate the clinical manifestations of intrusive events in actual patients and demonstrate the inherent complexities.

Posttraumatic Stress Disorder

From an index traumatic event, patients with PTSD typically experience two or three different, highly vivid intrusive memories both in visual and other sensory modalities (Grey and Holmes 2008). For example, after a traumatic road accident, a person might experience vivid intrusive visual mental images of an oncoming red truck, which originated from the moment just before the accident. This intrusive image is highly distressing and often associated with the strong emotions that occurred at the time of the trauma (fear, helplessness). Additional multimodal sensory images may originate from the same event. For example, the visual memory is also multimodal, comprising sight of the person's hand on the driving wheel accompanied by the sound of glass breaking and the smell of burning. The memory may also be associated with secondary emotions such as horror and guilt. Well after the event has occurred, an otherwise innocuous visual stimulus, such as a red front door, may remind the person of the red truck, thus triggering an emotional response. Emotional states may also serve as triggers for the traumatic event: when a person experiences a feeling of helplessness about an unrelated event, this may elicit the feelings of helplessness associated with the traumatic memory. These experiences may be especially disturbing because the patient has not associated these external or internal cues to the occurrence of the intrusive memory, so that they occur as if without warning and as both unpredictable and uncontrollable. Thus, even though these intrusive memory experiences

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

(colloquially often referred to as "flashbacks") may be brief (i.e., seconds, followed by up to 30 min of emotional response), they can have a powerful impact on ongoing behaviors and disrupt attention and performance (Holmes et al. 2017). Intrusive memories not only interfere with normal activities, they also produce significant emotional distress and physiological arousal which can be highly disruptive. The experiences typically occur repeatedly, at unpredictable intervals and in unexpected settings (i.e., elicited by various contextual stimuli), and may occur at varying intervals (e.g., one per week to several per day). Efforts to "push the memory" from one's mind often fail. The unpredictable nature of the events can set off a cascade of other symptoms, including efforts to avoid reminders of the trauma which may lead to social withdrawal. A primary goal for treatment is to reduce the frequency and emotional valence of the intrusive image-based memories.

There are several treatments with a good evidence base for PTSD. Behavioral treatments provide an interesting example from the perspective of intrusive memory. Trauma-focused CBT (cognitive behavioral therapy) and EMDR (eye movement desensitization and reprocessing therapy) both involve repeated exposure to the trauma memory. For example, trauma-focused CBT, a particular form of CBT, focuses on the trauma memory but not on verbal thoughts "about" the trauma (as CBT for depression would). It does this by using imaginal or *in vivo* exposure to the trauma memory, which requires the patient to bring to mind the sensory image-based memory in rich detail. For instance, over a series of 12 sessions, the patient is encouraged to talk about the trauma in detail, to "relive" the memory in their mind's eye, and when possible to bring in adaptive information for memory updating (e.g., feeling of safety, that they did not die in the car crash). Initially, this process is typically highly emotional (and the patient may become upset and cry in reliving sessions), but it becomes less so over repeated sessions. Patients are taught not to "avoid" reminders or to push the intrusive memory from their mind. In summary, the emphasis is on deliberately retrieving the emotional memory in vivid sensory detail. Over time the memory becomes less vivid, the emotions and meaning updated, and the number of intrusive memories declines.

EMDR is a related behavioral treatment that requires somewhat less detailed recall of the trauma: when the memory is recalled in a therapy session, it is done so in the presence of a concurrent task, such as side-to-side eye movements or bilateral beeps. The patient deliberately brings the trauma to mind and simultaneously performs the task guided by the therapist. Research indicates that the success of EMDR is related to impact of the concurrent task in working memory. This is not unlike the brief procedure being developed to reduce trauma intrusions by using a memory reminder plus *Tetris* computer game play (Holmes et al. 2009, 2010; James et al. 2015; Horsch et al. 2017; Iyadurai et al. 2018; Kessler et al. 2020; see also section below on Future Interventions).

At the end of successful behavioral treatment, the patient should be able to recall the traumatic event at will, if they wish to, without becoming

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

overwhelmed. Critically, they will no longer be experiencing frequent involuntary, unwanted intrusive imagery-based memories that impair their daily life. If you have not experienced intrusive events yourself, it may be hard to imagine how powerful brief intrusive events can be for an individual, and thus how beneficial it can be to ameliorate them. After trauma, intrusive thoughts can carry damaging and toxic meanings for the patient. For example, a rape victim may have an image of the rapist telling them they are worthless. Rationally, a victim may know this is not true but still suffer under extreme distress and shame, brought about by the intrusive events, which cause them to relive this toxic message vividly. Verbal discussion does not change the meaning carried by the emotional image (presumably because it differs neurally), but strategies to change the intrusive event (the image itself) can, as illustrated here.

It is clinically compelling to see how successful PTSD treatments are centered on ameliorating the "hub" symptoms of intrusive sensory memories of the traumatic event. Reducing the emotional efficacy of intrusive events and the meanings they carry, as well as the frequency of their occurrence, can lead to substantial improvements in a patient's quality of life.

Many questions remain: How can models explain the impact of existing treatments on the reduction in frequency of intrusive memories? How can treatments be improved to make them briefer and even more effective? Ideally, we should develop simpler, focused treatments that can help more people globally (Holmes et al. 2014, 2018).

Obsessive-Compulsive Disorder

In OCD, intrusiveness coincides with the feeling of "being out of control" and is experienced in different phenomenological domains, as OCD develops over time. OCD is a process with different clinical stages rather than one single stage. These stages follow a dialectic interaction in which an intrusive event elicits a response and the response amplifies the intrusive event. Through different neurobiological adaptations, the course of OCD eventually worsens. Thus, OCD should be regarded as a disease *process* that develops through the amplifying *interaction* between (the reflection and resistance of) the person (mind) and the disorder (brain).

Take, for instance, a young mother who recently gave birth to her first child. She carries an enormous burden of being solely responsible for a helpless and vulnerable life. Her husband leaves daily for work, leaving her alone at home with the child. Seeing her young baby in the crib, a thought appears in her mind: she imagines that she could strangle her baby in the crib and that because she is alone, no one could prevent her from acting on that thought. In summary:

1. The mere *presence* of this thought is intrusive because it occurs against her will. She feels out of control and is unable to volitionally control her thinking.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

2. She is worried because the idea of strangling her baby does not match the ideal of motherhood she maintains and strives to achieve. The *content* of the thought is intrusive because it is not in line with her identity or expectancy. It is *ego-dystonic* in that it does not match her self-image, thereby giving rise to mental discomfort and distress.
3. She wonders whether she really could strangle her baby. How can she be certain, given the fact that humans are notoriously unpredictable, that she won't in fact destroy that which she most treasures? It seems that the freedom to commit such a terrible act is in itself so disturbing that it causes the thought repeatedly to reoccur. In addition, *moral implications* of the initial thought intrude as well.
4. She feels anxious because the thought confronts her with feelings of being out of control. The *emotional value* (anxiety) of the thought is intrusive. *The presence, content, implication, and emotional value of the thought all have an intrusive quality.* Although she actively resists the thought because it annoys her and feels intrusive, the very process of reflecting and/or resisting the thought *reinforces the frequency and strength* of the intrusive thought. Her efforts to eliminate the thought may, in fact, enhance its occurrence.
5. The thought becomes obsessional, and her attention is completely drawn to that one single thought. Obsessionality is a dysfunction of *intentionality*: the incapacity to shift focus or attention to another topic, due to a stronger and longer intentional relation with the mental act. The thought is intrusive because of its *obsessive* nature.
6. She cannot suppress the thought; moreover, she is compelled to think about her obsession. Compulsivity is a dysfunction of *sense of agency*: she is forced to think about the intrusion, contrary to her willpower. The thought is intrusive because of its *compulsive* nature.
7. Gradually the thought becomes more present and repetitive; it loses its original meaning, but remains an intrusion because of its duration and repetition. The thought is intrusive because of its new form or *appearance*; it is now a full-blown *obsession*.
8. Obsessions are answered with compulsions. (Note: both obsessions and compulsions are intrusive, with both having obsessional and compulsive qualities.) Though initially successful in reducing anxiety, these compulsions gradually become intrusive since the acts have to be performed *compulsively*.
9. Eventually, the anticipatory power of the intrusion becomes so overwhelming that reality testing is disturbed. She does not know anymore whether she has or has not strangled her baby. Thoughts may become *delusion-like*, with psychotic features.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

Assessment and Domains of Intrusive Events

Because intrusive thoughts occur in healthy individuals as well as pathological states, it is important to develop sensitive tools to quantify these experiences. Healthy individuals experience unwanted intrusive thoughts (e.g., of dirt, contamination, doubt, harm, injury, sex, religion, order, symmetry, superstition) that are structurally and content-wise similar to clinical obsessions in OCD patients (Freeston et al. 1991; Langlois et al. 2000a, b) but less severe and disruptive. Intrusions vary in both the structure and content across individuals (Clark and Inozu 2014), as well as in terms of frequency, intensity (or distress), the degree to which the event is being perceived as an intrusion, unexpectedness, persistence (duration), controllability, vividness, valence (positive vs. negative), adhesiveness (durability), and modality (verbal vs. imagery based). These features provide dimensions that can be used to quantify intrusive events and to relate them to the underlying neurobiology. Thus, a refined multidimensional quantitative assessment of an intrusive event is critical for developing quantitative models and to assess the efficacy of interventions.

As reviewed by Clark and Purdon (1995), a number of investigators have developed questionnaires to assess intrusive events. These questionnaires distinguish between intrusive events that appear to be triggered by external stimuli and those that occur spontaneously. In one analysis, it was estimated that approximately 80% of intrusive events are provoked by an external trigger (e.g., Edwards and Dickerson 1987a). Clark and Inozu (2014) note that whereas intrusive events in nonclinical samples are context dependent (i.e., have external precipitants), clinical obsessions in patient populations appear to be more spontaneous. They also claim that avoidance of triggers in patients is associated with more adverse impact of both intrusive events and obsessions. Clark and Radomsky (2014) identify several aspects that need to be considered in assessing intrusive events:

- Content and process characteristics (the degree to which the event is unwanted)
- Discriminant validity
- Relationship to measures of worry and/or rumination
- Degree of self-relevance
- Appraisal variables (e.g., controllability, unacceptability, discomfort, guilt, dismissibility, unpleasantness)
- Degree of personal responsibility

Investigations of intrusions have taken several approaches, including questionnaires, diaries, and procedures that assess the impact of intentional mental control on unwanted intrusive thoughts. Each approach has limitations. The most common approach has been self-report questionnaires:

- The *Experience of Intrusions Scale* is a five-item measure that assesses the frequency, unpredictability, and unwantedness of intrusive thoughts,

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

as well as the interference and distress caused by the intrusions, each on a five-point Likert-type scale (Salters-Pedneault et al. 2009).

- The *Interpretation of Intrusions Inventory* consists of 31 items that refer to interpretations of intrusions that have occurred recently. Three of the above domains are represented: importance of thoughts, control of thoughts, and responsibility (Obsessive Compulsive Cognitions Working Group 2001).
- The *Obsessive Beliefs Questionnaire* focuses primarily on obsessive intrusion (Obsessive Compulsive Cognitions Working Group 2003).
- The *Obsessive Intrusions Inventory* consists of a 52-item self-report instrument designed to assess intrusive thoughts, images, and impulses that are similar to the aggressive, sexual, and disease-related thinking characteristic of clinical obsessions (Purdon and Clark 1993).
- The *Cognitive Intrusion Questionnaire* examines the following domains: frequency, duration, percentage of verbal and image content, interference, ego-dystonic nature, stimuli awareness, and associated emotions (Langlois et al. 2000a).
- The *Cognitive Intrusions Questionnaire–Transdiagnostic Version* items are grouped based on theoretical criteria into categories labeled intrusiveness, appraisals, emotions, and strategies, which were selected as components of a model that encompasses the different ways in which intrusive thoughts are processed (Romero-Sanchiz et al. 2017).

Whereas each of these questionnaires has a place in a specific context, they are limited as measures of intrusive events across diagnostic categories. Further, they were not designed with the goal of investigating the neurobiological basis of intrusive events. Some of the questionnaires include a broad range of negative thought content (e.g., anxiety and depressive thoughts) which are not, strictly speaking, intrusive events. Questionnaires are also limited because intrusions are often idiosyncratic and triggered by external cues that are difficult to describe: their dependence on retrospective self-report of unwanted intrusions may not be fully reliable. The questionnaires have not been used in the context of transdiagnostic, dimensional psychopathology, which are necessary to determine the heterogeneity of these events. In addition, they have not been designed to connect to the underlying neuroscience (i.e., the use of domains that can be mapped to specific brain systems), and the item response characteristics have not been rigorously assessed. Thus, future item bank-based questionnaires might substantially reduce subject burden by using an adaptive measurement framework similar to the PROMIS system (Cella et al. 2010).

An alternative to questionnaire-based approaches is the use of patient diaries. Clinicians using behavioral therapies for intrusive events typically ask patients to monitor their intrusions in a diary (Grey and Holmes 2008). In experimental studies with healthy volunteers, a range of diary measures for intrusive events have been developed so that participants can report on their intrusive

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

experience in daily life (James et al. 2016c; Lau-Zhu et al. 2019). Recording modes can range from pen and paper, to an SMS or an online interface. Diaries have the merits of being able to be precisely tailored to the research question in mind; they are sampled in real time and are thus less prone to memory biases inherent in retrospective self-report. One could imagine that apps will become useful in this regard.

Although intrusive events may be a component of worry, simultaneous administration of worry and intrusive event questionnaires yielded different factor structures. Nevertheless, the factor structure for the strategies used to counter the thoughts were highly similar for both types of thought. Furthermore, regression analysis identified interesting relationships between the strategies, the thought characteristics, and appraisal of intrusive thoughts and worry (Langlois et al. 2000a, b).

It appears that concern about the personal meaning of the thought is a unique dimension for obsessive intrusive thoughts (Clark and Claybourn 1997). There is, however, fundamental disagreement as to whether clinical and nonclinical intrusive events can be conceptualized along a continuum. Some researchers focus on thought content: Belloch et al. (2004) found that the ten most frequently occurring thoughts were related to accident, harm, sex, and aggression. Others focus on the process characteristics, thus emphasizing the intrusive aspect of the thought (Rachman and de Silva 1978). Interestingly, Lee et al. (2005) found that the most upsetting intrusive thought is often autogenous; that is, such intrusions come abruptly into consciousness without identifiable evoking stimuli and are perceived as ego-dystonic, aversive enough to be repelled, and include sexual, aggressive, and immoral thoughts or impulses.

Intrusive thoughts are thought to be closely related to dysfunctional beliefs (Obsessive Compulsive Cognitions Working Group 2003). Thus, there is an urgent need to assess underlying belief systems associated with intrusive events. These include:

- Over-importance of thought: beliefs that the mere occurrence of an intrusive thought marks its significance.
- Need to control thoughts: beliefs that one can and should exercise complete control over unwanted intrusive thoughts, images, and impulses.
- Perfectionism: beliefs that a perfect response or solution to every problem is necessary and that even a minor mistake can lead to serious consequences.
- Inflated responsibility: beliefs that one is liable for causing and/or preventing significant negative outcomes for self or others.
- Overestimated threat: beliefs involving exaggerated estimates of the probability and/or severity of harm to self or others.
- Intolerance of uncertainty: beliefs that it is necessary to be certain and that unpredictability and ambiguity should be minimized as much as possible.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

Another key aspect of intrusive thoughts relates to an appraisal, or how the intrusive events relate to important goals, values, and concerns. Appraisals follow intrusive thoughts and have been linked to future increased frequency. Appraisal can affect the ability to dismiss the thought, guilt, uncontrollability, and belief that the thought could come true. It could also affect responsibility, perceived consequences of the thought, beliefs about the importance of the thought, and worry that the thought may reflect something about one's personality (Berry and Laskey 2012). Moreover, in response to intrusive thoughts, individuals frequently engage in the following strategies:

- Reasoning that focuses on the thought being irrational or unimportant
- Thought replacement geared to distract or stop the worrying
- Social support, talking through, reassurance seeking, physical action, or doing nothing

While the themes of intrusive thinking are similar across clinical and nonclinical populations, the appraisal strategies in clinical populations seem to focus on responsibility and subsequent avoidance strategies. Results indicate that the more distressing a thought was perceived to be, the more likely participants were to recommend unhelpful strategies (Bomyea and Lang 2016). Conversely, the less distressing an intrusive thought was, the more likely participants were to recommend helpful strategies (Levine and Warman 2016).

Consistent with the appraisal model, Purdon and Clark (1994) found that the belief that one could act on the intrusive thought and a perceived uncontrollability of the thought were important predictors of the frequency or persistence of the distressing intrusion. Freeston et al. (1991) note three distinctive, dominant response styles: (a) no-effort response (26%), (b) attentive thinking (34%), and (c) escape or avoidance (40%). Moreover, they found that intrusions eliciting escape-avoidance strategies were evaluated more disapprovingly than thoughts eliciting attentive thinking. Rachman (2014) pointed out, however, that the following issues still need to be addressed:

- Additional information is needed on prevalence in clinical/nonclinical samples.
- The variable content of intrusions needs to be examined as a function of culture and environment.
- The nature and effect of repugnant versus nonrepugnant intrusions require examination.
- Further research is necessary on intrusive images and other percepts.
- Experimental investigations of intrusions are needed.
- Randomized clinical trials are needed to examine the effect of different treatments on intrusions.

Taken together, it will be important to assess appraisal domains and to quantify response styles associated with intrusive events.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

Following work and suggestions by others (summarized above), one approach to measure and quantify intrusive events is to decompose the experience into different domains or dimensions. For example, an intrusive event brought on by a ruminative thought that refers to a past experience associated with a strong negative emotion could be quantified along the following dimensions:

1. The degree to which the intrusive event (e.g., a short utterance or an elaborate verbal instruction) is characterized by a verbal thought
2. The degree to which the intrusive event (e.g., a vivid image, smell, sound, or other internal perceptual experience) is characterized by an internal sensory representation in the absence of a percept
3. The degree to which the intrusive event is associated with a positively or negatively valenced affect (e.g., severe anxiety, guilt, or shame), and the quality of that affect
4. The degree to which the intrusive event refers to an experience in the past, present, or is focused on possible future events

This approach would allow us to characterize the degree to which an intrusive event recalls a distant or recent past event or refers to an immediate or remote future event. For instance, the intrusive image experienced by an individual with PTSD might be characterized as relatively low intensity on Pt. 1, high intensity on the sensory representation (Pt. 2), with an association of guilt or shame (Pt. 3), and related to the past (Pt. 4). In comparison, an intrusive thought (worry) in an individual with generalized anxiety disorder might rate high intensity on Pt. 1, low intensity on Pt. 2, with a focus on anxiety (Pt. 3), and a primary focus on the future (Pt. 4). Such a decomposition could be used in large-scale surveys to begin to delineate the frequency of the phenomenology of intrusive events and the association of these events with a particular disorder. Subsequent statistical analyses (e.g., latent variable analyses and cluster analyses) could then be used to develop an empirically derived taxonomy of intrusive events. This approach would permit the severity of the intrusive event to be quantified and could be an outcome measure for the success of intervention studies.

Finally, this type of decomposition lends itself as a covariate for neuroimaging studies to delineate the neural circuitry associated with intrusive events. For instance, it would be extremely interesting to determine whether individuals who have experienced predominantly visual sensory intrusive events show changes in activation patterns in the visual processing stream, such as the occipital cortex, compared to individuals with similarly intense intrusive verbal thoughts hypothesized to show primary changes in the left ventrolateral prefrontal cortex.

How could these domains be neurologically implemented? In Appendix 17.1, we present a model that decomposes intrusive events into domains and then considers how these domains might be instantiated in the brain.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

Treatment Approaches

To date, most existing interventions consist of treatments that address each psychiatric disorder as a whole, rather than intrusive events. One notable exception is an innovative behavioral treatment for bipolar disorder, which specifically targets mental imagery-based intrusions through imagery-based cognitive therapy (Holmes et al. 2019). This same treatment has shown promise in reducing overall bipolar mood instability (Di Simplicio et al. 2016; Holmes et al. 2016a). In addition, a preventive intervention has been developed for PTSD that directly targets intrusive image-based memories of an experienced trauma (Iyadurai et al. 2018). The development of treatments follows three primary modalities:

First, *pharmacological treatments* exist for most psychiatric disorders that involve intrusive thoughts, although their efficacy varies widely across individuals and stages of the disease process. Selective serotonin reuptake inhibitors (SSRIs) are the first line of treatment for a broad range of disorders, from OCD to eating disorders, major depressive disorder, and PTSD. Patients who do not respond to SSRIs may be treated with serotonin and norepinephrine reuptake inhibitors (SNRIs), mood stabilizers, or anticonvulsants. Beyond these classes of drugs, there are numerous experimental treatments (e.g., hallucinogens, ketamine, atypical antipsychotics, and opioid drugs). We note that both SSRIs and hallucinogens, such as psilocybin, DMT and LSD, operate primarily on the serotinergic pathway: SSRIs increase the presence of serotonin in the synapse, whereas hallucinogens bind to the 5HT-2a receptor as a serotonin agonist. In both cases, a postsynaptic neuron will respond as if more serotonin is present. Some neural network modeling has suggested that an increase of serotonin in a network can facilitate the dynamic firing patterns of that network from getting stuck in local minima of processing behavior. There is a growing popular movement, albeit with limited empirical evidence, of using low doses (i.e., microdosing) of psilocybin and other hallucinogens to treat mood disorders, as well as high doses of psilocybin and ayahuasca to treat depression (Pollan 2018). For SUDs, a different class of pharmacological treatments exists: drugs used to treat SUDs are typically specific to the class of abused drug. For example, transdermal or oral nicotine may be used for smoking cessation, and opioid partial agonists are used for opioid disorder. Generally, these drugs take the form of agonists, partial agonists, or antagonists that target the receptor system where the drug acts. The idea is that the drug alters the receptor function so that the drug-related stimulus loses its incentive value. The intrusive event in SUDs often takes the form of strong cravings or urges to use a drug. Thus, drugs under development to treat substance abuse address these cravings as well as the actual drug use.

Second, *behavioral interventions* include a broad range of procedures: CBT, whose techniques include *exposure* to salient stimuli and reminders,

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

cognitive restructuring, as well as mindfulness and contingency management for SUDs, among others. Typical treatment targets are (a) to lessen reactivity to triggers that initiate the intrusive event or to update the underlying memory (e.g., after trauma) and (b) to develop strategies to control the user's response to the intrusive event. Treatments used to lessen the initial impact and/or update the memory include exposure therapies or repeated experiences with the thoughts or images in a safe environment. This may be seen as akin to extinction procedures in operant conditioning. A new form of treatment technique is being developed by Holmes and colleagues that targets the memory aspect of an intrusive event by interfering with consolidation or reconsolidation of the memory of a traumatic event (Holmes et al. 2009, 2010; James et al. 2015; Horsch et al. 2017; Iyadurai et al. 2018; Kessler et al. 2020). Treatments used to control the responses include cognitively reframing the meaning of the eliciting event, developing competing behavioral strategies that are incompatible with the immediate response, utilizing social support to minimize the emotional response, or learning to decrease emotional reactivity. For example, mindfulness training has been shown specifically to decrease habitual reactivity to intrusive events (e.g., craving for cigarettes and food) and has led in some cases to significant reductions in unwanted behaviors: five times the quit rates in smoking cessation and 40% reduction in craving-related eating (Elwafi et al. 2013; Brewer and Pbert 2015; Brewer et al. 2018; Garrison et al. 2018; Mason et al. 2018).

Third, *nonpharmacological brain interventions* include neuromodulation such as transcranial magnetic stimulation (TMS), deep brain stimulation (DBS), lesions to specific brain areas (e.g., anterior cingulate cortex), and electroconvulsive shock. These techniques have been used under limited circumstances and in highly selected patient populations. DBS, for example, is used only in severe cases of OCD (Tyagi et al. 2019). DBS targeting the ventral limb of the internal capsule and the nucleus accumbens is an effective treatment strategy for treatment-refractory OCD. TMS is approved to treat depression (Mutz et al. 2018) and is under study for the treatment of OCD (Rapinesi et al. 2019), Tourette syndrome, PTSD (Kozel et al. 2019), and SUDs (Zhang et al. 2019).

In many cases, pharmacological, behavioral, and neuromodulatory modalities are used in combination in various forms. Drugs are used to facilitate the psychotherapeutic process, such as MDMA and psychotherapy for PTSD (Mithoefer et al. 2019), whereas behavioral treatments are used in combination with nicotine replacement therapy for smoking. These different treatments may occur simultaneously or sequentially. In one case, the FDA has approved a treatment for depression using all three modalities: TMS, SSRI, and behavioral treatment. As noted, we have no information about the effects of these treatments specifically on the frequency or severity of intrusive events, separate from other symptoms of the disorder, and future research is needed here.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

Future Interventions for Psychiatric Disorders

As discussed, very few treatments exist that explicitly target intrusive events as an aspect of a disease, and for those that do exist, it is not clear how they actually modify intrusive events. Although intrusive events are included in the DSM-5 criteria for many disorders, they have not yet been targeted as a distinct treatment domain.

Treatment innovation is essential and may benefit from being mechanistically driven and by combining treatment modalities, such as combining pharmacology with a psychological/behavioral approach (Holmes et al. 2018). Before going into detail about specific treatment modalities, we discuss how interventions that target intrusive events might be used to modify core aspects of an intrusive event: gating, error correction, salience, and evaluation.

Gating Errors

Within psychiatric diseases, intrusive events may result from too much information being allowed into awareness, either in a global sense or by a selective memory gaining access to awareness through a selective gate. This broad construct could be applied to many psychiatric diseases. Intrusive verbal thoughts associated with general anxiety disorder (see earlier discussion) may be due to a "weak" gate that allows a lot of information in and ultimately leads to stress and anxiety. Intrusive image-based thoughts associated with PTSD may be due to a faulty assignment of salience and/or evaluation that emerges for the traumatic event, which allows them to enter into or persist in awareness more easily than nontrauma-related thoughts. Intrusive thoughts of drug cues in SUD may be construed as a combination of faulty gating and salience, compounded by reinforcement learning.

Error Signal/Detection

Intrusive events may also result from a heightened error signal during otherwise normal thought. To achieve ordinary activities of daily living, mental processes must stay on course and not be derailed by irrelevant streams of information that constantly come into our brain. A lot of information is processed on a subconscious level, wherein only the stimuli that are most different from our expectations are processed consciously. For example, when walking down the street, an individual is able to maintain posture and balance, typically via subconscious processing that has become highly automatized. Only when an event occurs that violates expectations (e.g., when we stumble or suddenly see an unexpected object in our path) is the walking process brought to the level of consciousness via bottom-up mechanisms. Alternatively, our own cascade of ongoing neural processing related to other aspects of life, which are not directly associated with the goal, can also interrupt automatic behaviors, as when we

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

suddenly remember that we were supposed to pick up the kids at school. Thus, bottom-up and top-down internal interrupts exist as well as external interrupts. The difference between our expectations of internal and external stimuli and the actual perception of these stimuli can be considered a prediction error.

It is possible, for example, that verbal or visual intrusions in PTSD may be related to a momentary yet large prediction error. That is, at times, the perceived difference between the thoughts associated with the trauma and the typical subconscious narrative that is ongoing in an individual's mind is large (high prediction error) and triggers the individual to switch their attention to those momentary thoughts. This is a particularly attractive hypothesis given the sparse temporal nature of intrusive events in these patients. In daily living, for instance, many individuals with PTSD are able to function relatively well in between intrusive events of their trauma. They are often able to work and care for families, and to conduct their lives for some days without having an intrusive event. The frequency of occurrence of an intrusive memory of trauma for some patients may be infrequent (e.g., once a fortnight) whereas for others with more severe levels of PTSD, it may be up to every hour. Typically, measurement tools of PTSD capture this rate of occurrence, although more research is needed.

Here we address emerging ideas for future treatment innovation based on the model proposed in the Appendix. Assuming that these reward prediction errors are coded in a specific neural network, a technology capable of sensing the magnitude of the violation between expectations and actual input may be able to change activity in this network in a dynamic manner, pushing it back into the intended state. This type of closed-loop neuromodulation is currently used in the treatment of epilepsy. Briefly, sensing and stimulating electrodes are placed in the brain of individuals with intractable epilepsy in the vicinity of the seizure focus. The device has a certain "tolerance" for background variability in the neural activity in the vicinity of the electrode. Once a critical level of variance is detected (high variance), the device is able to stimulate the brain and push it back into a healthy state (low variance), thus avoiding the cascade of a seizure (for further discussion, see Hanlon and McTeague, this volume). It is easy to see how a similar approach could be used to abolish intrusive events in a dynamic manner in individuals with PTSD. In veterans with PTSD, for example, a device could be trained to identify the "background" levels of activity present in normal life as well as activity associated with intrusive war-related memories, and permit the device to push the system back into the healthy range of error. One problem with this approach, however, is that large prediction errors are a crucial element for flexible human behavior. So, an autonomous closed-loop system would have to be able to detect differences specific to the trauma memory. While this seems like a tall order, because these intrusive events are so debilitating to the patients, it is reasonable to imagine that the amplitude of their prediction error is so large that the device would be able to have a very high tolerance threshold, and thus only produce stimulation during the

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

most extreme examples of prediction errors. Alternatively, individuals could be trained to self-administer the stimulation when they start to become aware of an intrusive event. More basic neuroscience research with a fast temporal sampling profile (e.g., EEG, MEG, *in vivo* recording) is necessary to evaluate if prediction error is a viable treatment.

Evaluation

One of the most widely implicated behavioral domains thought to be responsible for the maintenance of intrusive events is an aberrant evaluation system. Evaluation brings together goal hierarchy (including the current task, homeostatic goals, etc.), salience, affective, and reinforcement properties of previously learned behaviors to determine whether to stay on a certain task or switch. In terms of intrusive events, which can be seen as an alternate task, evaluation would help to determine whether to switch, and for how long, to the new task. Clinically relevant downstream manifestations of the evaluation system include frequency and duration of intrusive events in consciousness. In light of emerging understanding with regard to intrusive events, these manifestations can now be linked to specific neural systems, informed by what is known with current treatment paradigms. For example, treatments such as CBT target the cognitive elements of intrusive events, theoretically changing belief systems related to intrusive events (e.g., "Is this true?"), whereas mindfulness training targets the relational affective component of the intrusive event (e.g., "How caught up am I?"). These examples provide concrete treatment modalities that can be more critically studied with regard to efficacy. For example, does mindfulness training (as a modality to improve evaluative accuracy) change the "stickiness" of an intrusive event, and does this change the frequency, duration, and salience of the intrusive event in the future (see Brewer et al. 2013, 2018, 2019)? Further, the neural systems that are affected by these interventions can be more specifically studied. For instance, recent work with mindfulness training has linked reduction in default mode network activity with a reduction in cigarette smoking (Janes et al. 2019), yet the frequency and duration of intrusive events related to smoking have not yet been evaluated. Future studies can use experience sampling or other modalities to determine if intrusive events reduce in either of these domains.

A Neuroscience-Derived Neuromodulatory Approach to Treatment

Pharmacological treatments may change the sensitivity of the gate that allows items to enter into conscious awareness. For example, benzodiazepines may blunt emotional response to negative thoughts in a nonselective manner. Alternatively, there may be potential for combination treatments that couple a behavioral approach (e.g., one that evokes intrusive events and thoughts,

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

such as cue exposure), to open the gate, with a pharmacologic agent, to blunt the response.

Imagine if one could simply "zap" the brain and magically remove the intrusive event causing the distress. Such an idyllic scenario is (unfortunately) still far away, but promising research is moving us closer. A whole new field of treatment research has emerged around specific brain stimulation techniques due to their noninvasive nature and simple theoretical motivation. The central idea behind brain stimulation and modulation techniques takes a systems perspective, which considers the brain as the source of any behavior (the outcome of brain processes). Behavior is therefore the dependent variable whereas brain circuits, activity, or neurons are the independent variables. If we are to change a pathological behavior, this systems-level interpretation implies that we have to find, target, and modulate the brain process or instantiation that is generating this specific behavior.

Two main approaches have produced important results in recent years: TMS and neuroimaging-based neurofeedback. TMS is a technique that generates weak electrical currents within the brain through direct application of electromagnetic flux over the scalp, noninvasively into brain tissue (see also Hanlon and McTeague, this volume). At the electrophysiological level, TMS induces changes in neuronal excitability, which can cause changes in behavior. TMS has been recently approved by the FDA for use in the treatment of several psychiatric disorders, such as addiction and OCD. Future studies need to evaluate how TMS in combination with other therapeutic forms might be used specifically to resolve intrusive events.

In neuroimaging-based neurofeedback, "neurofeedback" is defined as a closed-loop procedure whereby online feedback of ongoing neural activity is given to the participant for the purpose of self-regulation. Noninvasive neurofeedback can be implemented with several neuroimaging modalities. Some of the most promising proof-of-concept cases derive from functional magnetic resonance imaging (fMRI) and electroencephalography (EEG)-based neurofeedback (Sitaram et al. 2016; Taschereau-Dumouchel et al. 2018a; Keynan et al. 2019). We will first consider fMRI, due to its potential as an acute intervention (used once or a few times).

With fMRI, one simple option is to use the overall signal strength in one predefined brain region. This may work well when the target is general (e.g., motor action initiation vs. no motor activity); in other cases, however, one needs to access a specific representation, which by definition would not be retrieved through the overall activation level. Therefore, rather than focusing on the overall activity level of a region (*unspecific*), machine-learning algorithms now allow us to infer the precise activity *pattern* that corresponds to a unique stimulus, object, or category (*specific*). This approach, borrowed from machine learning, is termed multivoxel pattern analysis (MVPA) (Kamitani and Tong 2005; Norman et al. 2006). Going even further, methods have now reached

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

the point where we can infer the pattern of brain activity in a target participant from brain activity patterns in surrogate participants (Haxby et al. 2011).

Although these advances have often remained outside the realm of clinical applications, recent innovative efforts have been directed toward utilizing MVPA in the field of neuropsychiatric disorders. While mainly centered around the development of markers for certain disorders or their subtypes, with some remarkable results (e.g., Yahata et al. 2016; Etkin et al. 2019), some studies have explicitly targeted the prediction of intrusive events from fMRI activity patterns (see Holmes et al., this volume; Clark et al. 2014b, 2016).

Bringing MVPA to a real-time setting leads us to the concept of neurofeedback, which essentially monitors brain activity patterns over time while using the machine-learning prediction as neuromodulatory input (e.g., to provide rewards). If an algorithm is able to detect intrusive events, the system may then provide a reward to remodel the association between brain state and appraisal or some other form of modulation to disrupt the cascade of neural events following the emergence of the intrusive event.

In terms of clinical development, due to the intrinsic nature of MVPA-based designs, neurofeedback interventions can be easily used in a double-blind, placebo-controlled way. Since we can find activation patterns that correspond to specific, distinct mental representations, we can also let the algorithm randomly choose which representation will be the target and which the control. Neither the experimenter nor the patient has to be aware of the category for the software procedure to be deployed effectively.

Research has shown that neurofeedback may work via reward processing, learning, and control networks (Sitaram et al. 2016), and that it depends on reinforcement learning processes (Shibata et al. 2018). During a typical experiment, the machine-learning algorithm monitors activity in a selected brain region and, at predefined time intervals, computes a (monetary) score reflecting the likelihood that the current activity pattern resembles a template, target mental representation. Over time, the brain learns to associate the occurrence of such representation with rewards.

In two proof-of-concept studies, this technique was used to reduce fear responses in healthy individuals conditioned toward simple visual stimuli (Koizumi et al. 2016) as well as in participants with subclinical phobia toward their feared object (Taschereau-Dumouchel et al. 2018a). The physiological fear responses (amygdala reactivity and skin conductance) were diminished only for the targeted representation, while a control stimulus elicited unchanged fear responses. More recently, the same group reported on the feasibility of using decoded neurofeedback as a target treatment in PTSD (Chiba et al. 2019). Importantly, throughout these studies, participants were not told about the link between their brain activity and the amount of reward they received on a trial-by-trial basis. When asked to make a forced choice about the target and control categories, participants answered randomly (Shibata et al. 2018).

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

The aspects introduced so far raise one crucial point: the entire neurofeedback intervention can be applied without ever mentioning or displaying the upsetting content or event to the patient. The enormous advantage that this could bring is evident: reducing implicit, physiological fear responses could create the foundation for subsequent, and more effective, behavioral or cognitive therapies. Neurofeedback could then, as a second step, be applied for maximal effects and learning.

Alternatively, neurofeedback modulation may provide the basis for dampening a physiological, autonomic reaction, which may then facilitate subsequent cognitive or behavioral therapies. Current work on fMRI, MVPA-based neurofeedback can be expanded to include the notion of dynamic brain state. Indeed, utilizing a very simple correlation of activity fluctuation between two brain areas, or a pattern of activity representing an object or category, may be too simplistic, in particular if we are to target states that are more global (e.g., attention, emotions, arousal, interoception). Spatiotemporal oscillatory patterns describe relatively well these global states, distinguishing between rest and task-based efforts (Vidaurre et al. 2017; Bolton et al. 2018).

Rather than acting on the intrusive events themselves, dynamics-based neurofeedback interventions could target and modulate an overall affective or cognitive state. Here, the rationale is different: prepare the brain to receive the treatments targeting specific aspects of the pathology in the individual.

fMRI neurofeedback carries high costs, immobility, and the requirement for specialized operating personnel. This greatly hampers the scalability of this approach, particularly if we think about going beyond acute treatments.

As such, EEG-based neurofeedback holds potential from a different perspective. New generation EEG headsets have relatively low costs, are small in size, and are almost "plug-and-play" ready for use. We can envisage EEG products in the near future that could be used autonomously by patients on a daily basis and essentially without many constraints in terms of location or function, allowing for real scalability. Proof-of-concept EEG neurofeedback has been demonstrated with specific mental states while targeting deep brain structures, such as the posterior cingulate cortex which is part of the default mode network (van Lutterveld et al. 2017), training stress resilience through electrical fingerprint (Keynan et al. 2019), as well as reduction, consolidation, and personalization of lead placement (Pal et al. 2019).

A different form of feedback might therefore involve generating an external interrupt when an undesirable thought pattern has emerged. For example, MVPA of EEG signals might allow ruminative thoughts associated with depression to be decoded. If detected, an EEG headset could then be designed that might vibrate or emit a tone whenever rumination had exceeded, say, ten seconds, bringing the patient back to the present.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

Combination Treatments

A significant future direction in the treatment of intrusive events is the possibility of combining treatment modalities. The interventions used in the treatment of intrusive events fall broadly into three categories: behavioral, pharmacological, and neuromodulatory. Some of these treatments have already been combined. For example, MDMA is currently under investigation for the treatment of PTSD, but a central feature of the treatment is that the drug is administered in the context of psychotherapy. Other examples include the use of behavioral treatments combined with nicotine replacement therapy for smoking cessation or other pharmacological-behavioral combinations for treatment of other addictions.

An Innovative Approach to Targeting Intrusive Events in Trauma-Related Disorders

In PTSD, trauma is induced through vivid, emotionally laden intrusive memories that take the form of sensory multimodal mental images of specific moments experienced during the traumatic event. Current treatments are available to reduce these memories or their consequences (e.g., trauma-focused CBT), but they can be hard to implement on a large scale (to reach more people), and evidence-based preventive interventions after trauma are lacking. Recently, Holmes and colleagues developed a brief, noninvasive behavioral treatment that reduces both the establishment and the maintenance of intrusive memories after a traumatic event (Iyadurai et al. 2018). It should be noted that the treatment focuses on specific intrusive memories, rather than the whole disorder of PTSD. The type of intervention developed may be applied either shortly after the trauma, or later, well after the memories have been established (Holmes et al. 2009, 2010; James et al. 2015). The procedure consists of the following steps:

1. Subject is instructed to recollect the event briefly via images and thoughts.
2. Subject participates in a 15-minute *Tetris* task that involves mental rotation.

This simple procedure can be implemented shortly after the actual event or after a retrieval of the event later in the process. It may be administered in a hospital emergency room or acute ward within six hours of the traumatic event (Horsch et al. 2017; Iyadurai et al. 2018). In these studies, the intervention reduced the number of intrusive memories in the following week by approximately two-thirds. It can also be used to treat intrusive memories of traumatic events that occurred many years ago. In one study (Kessler et al. 2018), inpatients with complex PTSD received the treatment with intrusive memory, resulting in a reduction in the frequency of those intrusions (compared to non-targeted intrusions). Further research is required.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

This novel treatment approach is derived from cognitive psychology and experimental psychopathology related to memory processes. Performing a visually demanding task shortly after encoding or retrieving a remembered mental image, while holding the image in mind, competes for working memory resources in a way that interferes with consolidation/storage (during the initial experience) or reconsolidation (after the image is retrieved later). Remarkably, performing this simple *Tetris* task immediately after thinking about the image reduces the frequency of intrusive mental images (Iyadurai et al. 2018). The technique is based on various assumptions:

- Intrusive memories of trauma comprise sensory mental imagery (Grey and Holmes 2008).
- Intrusive memories can be altered shortly after an event or at retrieval: memory consolidation/reconsolidation (Visser et al. 2018).
- The capacity of people's working memory is limited (Baddeley 2003).
- Visuospatial tasks compete for resources in working memory with a mental image; that is, those that would be needed to (re)consolidate intrusive mental images (James et al. 2015).

Thus, engaging in a visuospatial task, such as a highly visually demanding computer game like *Tetris*, at a time when the mental image of the intrusive event is active may reduce the reoccurrence of distressing images later, as well as the level of distress and vividness associated with them. Further research is needed to develop this relatively simple and brief behavioral intervention approach. It lends itself to be studied as part of a combination treatment approach, as discussed above. The difference between this theoretical approach and many others is the focus on a single symptom (intrusive events) and a neuroscientific account about the underlying mechanisms bringing about effects.

Implications Beyond Neuroscience and Psychiatry

Philosophical and Social Implications of Definitions

Although the original definition by Clark (2005) and its minimum and maximum derivatives are intuitively very recognizable, the appearance of clarity is misleading. The problems that come to light reflect some important philosophical and social assumptions.

What exactly is meant by unwanted? Is it a mental event that is unwanted by the individuals themselves? Or does this represent what is unwanted by members of society, relative to what counts as unacceptable, abnormal, or even moral or immoral (e.g., expressions of sexuality or aggression)? Is what is unwanted not wanted relative to short-term individual goals, such as wanting to drink water, eat a marshmallow, or court a colleague? Or is what is unwanted

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

defined relative to interference with the fulfillment of long-term goals, such as completing a PhD program or maintaining a marriage? Are intrusions perhaps unwanted in light of the conflicts they present to a person's ideal image, such as becoming a good scientist? Certainly the notion of being unwanted depends on a broader philosophical perspective: What kind of human being do we have in mind as our ideal? Are mental phenomena and interventions that foster this ideal more wanted than those that lead away from this ideal? In sum, there are multiple ways to be unwanted, whether by the individual or society. Should we privilege one type of being unwanted over another?

Inherent in the definition of intrusive thinking is the notion that a lack of control over the flow of mental events is unwanted. The implication is that intrusive mental events are unwanted at least in part because they are not subject to volitional control. Overcoming a loss of control, in turn, implies that mental events could be made subject to volitional control. But what makes a mental or neural process volitional versus nonvolitional? In the motoric domain of eye movements, for instance, changes in pupillary size (or nystagmus or microsaccades) seem to proceed automatically, regardless of intentions or plans held in working memory or reportable by a subject. In contrast, saccades and smooth pursuit eye movements are subject to flexibly updateable plans or intentions held in working memory, which in turn can be reported by a subject (e.g., "I was looking around for my child"). Implicit in this is the notion that different intentions would lead to different eye movements. Similarly, volitional control of mental events implies that different mental events could and would have arisen had intentions or plans been different. There seems to be an implicit assumption here that events could have turned out otherwise than they did, had volitional control of actions or thoughts been different.

This raises the age-old problem of free will. If, under a deterministic worldview, mental events could not have turned out otherwise, then whether a particular intrusion would happen at a given time is dictated by the laws of physics before one was even born. The occurrence of involuntary intrusions seems consistent with the possibility that they could not have turned out otherwise. It is the notion that they can be brought under voluntary control that may be undermined by determinism. But how would indeterminism save the possibility of voluntary control of mental events? It might seem that randomness is as little subject to agentic control as events determined to happen before one was even born. Is there a middle path between the apparent lack of voluntary control that arises either under determinism or utter indeterminism? If yes, what might that solution be?

Our definitions of intrusion demand that we specify what it means to be clinically relevant. No absolute answer is possible, only one relative to a given clinical approach, of which there are many. Moreover, relevance depends not only on the clinical picture, but also on the person being treated, as well as the diagnostician and the social context or culture in which the patient and diagnostician find themselves. This implies that it is never possible to determine

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

with objective certainty what an intrusion means in the psychopathological sense. It is therefore impossible to find a single neurobiological substrate of an intrusion without first specifying with great precision what exactly one means by seemingly basic terms such as "unwanted," "control," "volition," or "clinically relevant."

This definition of an intrusion refers to the fact that an intrusive mental event should be distinct or temporally punctate. We may try to apply this definition to "mind wandering," which is sometimes considered an intrusive event. Mind wandering, however, is often difficult to distinguish from regular thought patterns or normal free associations. It is certainly not punctate, but rather durationally extended. Would we regard the transition to mind wandering as an intrusion?

Moreover, is an intrusion by definition negative? Not every intrusion is accompanied by a negative emotion. Some intrusions, such as love or manic thoughts, are experienced as pleasurable, or may be concurrently both wanted and unwanted. Further, some types of intrusion are wanted by the subject, but viewed as intrusive by others in society. For example, in a traditional Nepali village, where marriages are typically arranged, the statement "I have fallen in love" might be met with concern rather than happiness, whereas in the West this transition to a potentially obsessive mental state is commonly regarded as positive. Patients with bipolar disorder report experiencing highly positive mental imagery intrusions associated with their mania, but these may also be diagnostic for a major psychiatric disorder (Ivins et al. 2014). Interestingly, one study has demonstrated that positive intrusive thoughts can be induced even in healthy adults, suggesting that they are open to experimental manipulation and study (Davies et al. 2012).

Although our focus in this chapter has been on the neuroscientific basis of intrusive events, their clinical relevance, and interventions, consideration should also be given to the degree to which intrusive events affect our society or can be understood outside of a solely biological or medical perspective. Radomsky et al. (2014) found that intrusive events are experienced across a large variety of cultures. They concluded that there were far more similarities than differences across different cultural sites and that the contents centered worldwide on themes of contamination, aggression, doubt, blasphemy, immorality, sex, victimization, and miscellaneous intrusions. Culture seems to influence the content of intrusions to some extent but not its prevalence (Clark and Inozu 2014). Moreover, the relationship between appraisals, control strategies, and the frequency and distress of intrusive events appears invariant across countries. This has been confirmed by others, who conclude that there is a certain degree of universality regarding the prevalence of obsessive, dysmorphic, hypochondriacal, and eating-related intrusions across a variety of countries and cultural contexts (Pascual-Vera et al. 2019). Nonetheless, there are cultural differences in how intrusive events are interpreted. For example, according to Luhrmann et al. (2015), voices are experienced by schizophrenics in the United

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

States as primarily negative, in particular, as violations of thinking, whereas in India and Ghana they tend to be interpreted in a more positive light, involving relationships with the presumed speakers. Understanding intrusive events in different cultural contexts will constrain the degree of generalizability of any neuroscience-based model of intrusive events. Future research should examine individuals who report similarly frequent or intense types of intrusive events in different disorders, but who are members of different cultures and have varied responses to intrusive events. Such an approach is not unlike extracting different computational models from psychotic and nonpsychotic hallucinators (Powers et al. 2017). Taken together, intrusive events can be a fruitful and important topic of research that extends the biological sciences and can provide important information about how society should consider managing these phenomena in the future.

Of particular interest is the change in the nature of intrusions experienced by people in our modern societies. Those who came of age in the 1980s or earlier effectively grew up in an analog world. Since the 1990s, however, with the advent of personal computers, the Internet, and smartphones, our society has become digital. Because the last generation to be raised without digital devices is still alive, now may be an opportune time for this generation, which has experienced both the analog and the digital world, to reflect on the pros and cons of this societal and personal transformation.

Modern society is now deeply penetrated by mobile devices, which can be viewed as "intrusion machines." Screen media activity (SMA) is ubiquitous worldwide and among the most salient recreational activities of children and adolescents. Children and adolescents spend about 40–60% of their time after school engaged in SMA (Arundell et al. 2016), and nearly 97% of U.S. youth have at least one electronic item in their bedroom (Hale and Guan 2015). A heated debate has emerged on whether SMA is associated with psychological and social problems (Ferguson 2017; Twenge et al. 2017). However, media behavior is complex, encompassing a variety of activities, such as social and nonsocial Internet use, gaming, as well video or TV viewing. For example, whereas males are more likely to engage in video games with a higher potential for excessive use (Choi et al. 2015), females engage more with social media (Schou Andreassen et al. 2016) and exhibit more excessive cell phone use. Moreover, gaming has replaced sedentary screen time, such as TV viewing, Internet usage, and nonactive gaming (Simons et al. 2012). In the near future, with 5G technology, individuals may be able to interact with numerous devices on an almost constant basis. For instance, whereas currently a cell phone might signal the arrival of a text, email, or message (the intrusive event), we may in the future be alerted by our refrigerator, car, air-conditioning system, or other devices that we use in our daily life. Thus, technologically based intrusive events could have the potential to seriously affect healthy individuals and possibly to a greater extent any individual who is cognitively or affectively compromised (i.e., a person with a psychiatric disorder). Moreover, the constant

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

engagement with SMA may affect both brain structure and function, and in turn make individuals more susceptible to the experience of both physiological and pathological intrusive events. The investigation of mobile technology on cognition is still in its infancy (Wilmer et al. 2017), and our knowledge about the impact of this technology on intrusive events is nonexistent. Thus, an important goal for future studies should be to determine whether such devices could contribute to the exacerbation of psychiatric disorders, characterized by frequent and/or severe intrusive events.

Another perspective on intrusive events and their pathology is the notion that an individual who experiences a low level of need for control might not experience intrusive events as problematic or in need of treatment. This raises the question whether the fact that we are focusing on neuroscience, clinical consequences, and interventions associated with intrusive events is a by-product of our society's focus on "over control." Interestingly, individuals with strong beliefs about controlling thoughts are more likely to experience distressing intrusions, both with and without meta-awareness, compared to people with weaker beliefs (Takarangi et al. 2017). Thus, it would be interesting to examine the frequency, severity, and clinical consequence of intrusive events in different societies that place different emphases on cognitive control. Along similar lines, an intrusive event might only be assessed as an unwanted perturbation if the individual has a concept of causes and effects, which is a historical consequence of the Enlightenment period of the seventeenth and eighteenth centuries. In prior societies there was a greater emphasis on a teleological framework within which experiences and events were interpreted through the lens of their potential function, end, purpose, or goal. In this context, an intrusive event may be experienced as something that is necessary to lead to a particular goal rather than an unwanted distraction. Thus, it may be interesting to conduct a historical literature analysis that focuses on the characterization of intrusive events before and after the Enlightenment period.

Let us return now to intrusive events and tie them in with the notion of frameworks of meaning. If efforts to suppress intrusive events tend to exacerbate intrusive events, then perhaps efforts should be made concerning how best to suppress such suppression, or, on the contrary, how best to facilitate the expression of intrusive events so that the salient issue so expressed can be processed in a healthy manner. Other cultures have created modes for the expression of "forbidden" emotions. Ancient Greek drama often centered on creating tensions that would then lead to emotional catharsis. Rather than suppress unwanted emotions, such as lust for forbidden objects of desire, as in the play Oedipus Rex, these emotions were vented in a manner that was safe for society. Even modern European cultures have aspects that are reminiscent of catharsis. For example, Carnival is a venue for the expression of carnal desires and behaviors that in other times would be regarded as deviant or dangerous. How might catharsis be exploited in the context of existing or new therapeutic methods? One possibility would be role-playing, where a traumatic event

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

or idea is expressed in a manner such that emotions can be safely expressed. Indeed, by pretending to reenact a particular traumatic event or relationship, the possibility emerges of having it not only expressed, but of expressing it in a new way, perhaps having the event turn out differently than it in fact did. Another avenue might be to build on the tragedies and plays of Ancient Greece by creating "virtual worlds," perhaps exploiting movies or virtual reality technology, where pent-up and suppressed emotions and desires could be released virtually, rather than through real acts in the life of the patient.

According to this view, intrusive events are analogous to salience signals arising from exogenous attentional circuitry. Just as the sudden motion of a tiger demands an interruption of the current plan engaging conscious thought and planning, so that a new plan (in this case, to escape the tiger) can be generated, intrusive events are salience signals that enter consciousness because there is an unresolved issue that requires executive control circuitry to come up with a plan to resolve the unresolved issue. If this view has validity, then suppressing the salience signal might be about as effective as attempts to suppress hunger or thirst signals. The reason these signals barge into consciousness is so that goals can be generated by executive planning areas that will resolve them (e.g., coming up with a plan to get food or water). If unresolved emotional or cognitive issues barge into consciousness, a better approach than suppression may be to find ways to resolve the unresolved issue. Catharsis and role-playing have already been mentioned. Other possible methods may involve unorthodox techniques that are considered quite orthodox in non-Western traditions. For example, according to Kundalini yoga, emotions are stored in the body and can be activated with certain bodily actions, such as breathing patterns of physical exercises. By intentionally invoking the breathing pattern associated with fear or calmness, say, the mental state that normally accompanies such breathing patterns can also be invoked and perhaps processed in a manner subject to volitional control. If unresolved emotional issues are stored in part in the body, or in bodily patterns of action, perhaps the Western tradition can gain insights from other traditions concerning the "cleansing" of stored psychological tension and pain.

Conclusion

Intrusive events are emerging as an area of interest that can help further our collective understanding of basic brain function, psychiatric conditions, and their treatment. Characterizing domains in which intrusive events manifest in psychiatric disorders may help their characterization, diagnosis, and treatment. Basic heuristic models can inform how normal brain function can go awry due to faulty systems, including gating, salience, evaluation, and prediction error detection. From these models, current and evolving treatments (and potential combinations) can be tested for target engagement, specificity of effect

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

and efficacy with regard to reduction of frequency, duration, and salience of intrusive events. Future research would benefit from mechanistically framed and neuroscience-based approaches with specified targets and tangible target engagement and clinical outcomes. Our proposed model (see Appendix 17.1) provides an example of one such emerging approach of a behavioral intervention that specifically targets intrusive events based on theories of memory (re) consolidation and cognitive task interference.

Appendix 17.1: Proposed Model

Our model begins with a sequence of goals that are represented in the brain (Figure 17.A1). A mental workspace keeps the present goal in mind, allows operations to take place over representations held in the workspace (area 4, working memory), and takes into account prior knowledge, context, and other system constraints. Salience helps the organism determine how important a potential interrupt is and interacts with an evaluative system to determine whether to stay on task or to interrupt it and, if so, which task to then prioritize as the task to do next. A goal maintenance system helps the mental workspace maintain the present goal using feedback loops that afford the minimization of prediction error signals. This may happen via the enhancement of gating of potential interrupt inputs, or the inhibition of the salience of potential interrupt signals. The goal maintenance system evaluates deviations from the trajectory leading toward fulfillment of the goal. Such prediction error signals are used dynamically and cybernetically to correct the present trajectory to minimize that error, similar to when a heat-seeking missile alters its path to hit its target.

External interrupts (not internally generated) may orient the organism to unexpected inputs from the external world (e.g., when we hear a loud sound or see a sudden motion): if the magnitude of the prediction error (i.e., between what was expected and what in fact occurs) is large enough, the organism orients to the external stimulus.

Internal interrupts can be both bottom up and top down. Bottom-up systems that can generate interrupts include systems that maintain physiological (e.g., hydration) and nonphysiological (e.g., happiness) goals that are separate from the current goal. When salient enough, these interrupts provide inputs to the mental workspace to force a reprioritization of what to do next (stay on task or switch tasks). Subjectively these signals are experienced as, for example, thirst, hunger, lust, or a need for oxygen, salt, or sleep. Other bottom-up systems include reward/punishment and other evaluative systems but may not have intrinsic homeostatic functions. These may be experienced as, for example, fear or other emotions, such as anger. In addition, the memory systems may automatically retrieve memories which can then appear to "pop" into consciousness.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

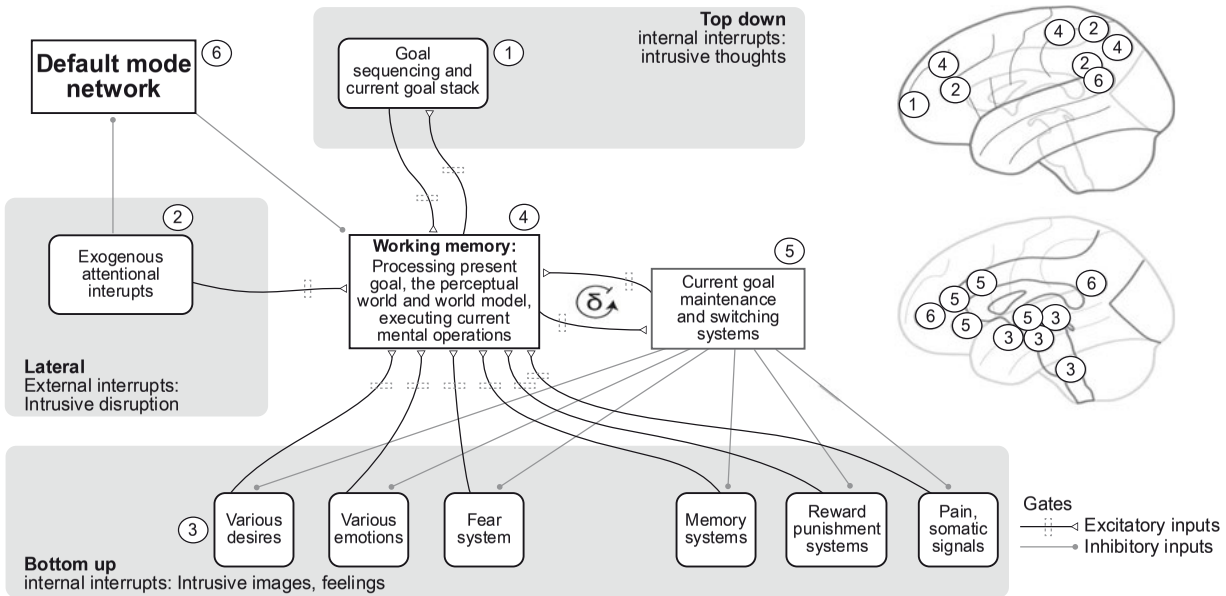

Figure 17.A1 (1) Neural substrates: frontopolar cortex (e.g., Brodmann area 10). (2) Ventral attentional network: ventrolateral prefrontal cortex, temporoparietal junction. (3) Numerous subcortical systems including amygdala, hippocampus, nucleus accumbens and ventral tegmental area, brain stem, and thalamus. (4) Dorsal attentional network: dorsolateral prefrontal cortex, posterior parietal cortex, superior parietal lobule, and intraparietal sulcus. (5) Cingulo-opercular control network: dorsal anterior cingulate, frontal operculum, basal ganglia; for emotional regulation ventromedial prefrontal cortex (BA 12) and orbitofrontal cortex (BA 11). (6) Default mode network: anterior medial prefrontal cortex, posterior cingulate, and temporoparietal junction.

Thus, under normal conditions, a goal is maintained using gating, evaluation, and prediction error monitoring and minimization. Under normal conditions, gating is appropriate, evaluation is accurate, and prediction error is accurate and dynamically corrected. Under pathological conditions, gating can become too weak or too strong, and prediction error and evaluation can become inaccurate. This model suggests that any one of these processes, or a combination of gating, evaluation, and prediction error, can go awry and thereby cause a pathological condition.

Failure Modes

Pathological conditions may arise when an individual fails to stay on target in the pursuit of the current goal because

- the prediction error (Figure 17.A1, between areas 4 and 5) does not arise (e.g., one should feel guilt for a misdeed, but does not) or it is inaccurate,
- evaluation (i.e., signals generated from the different components of Figure 17.A1, area 3) becomes inaccurate, or
- prediction errors do not get eliminated after a course correction takes place.

A system can have a number of modes of failure when one attempts to fulfill one's goals (see Table 17.A1). For instance, in a normal state, if the current goal is to go to a movie, the following thought might arise as a bottom-up interrupt: "You didn't study enough for your exam next week!" If gating is appropriate, that thought need not get into the mental workspace. If the gate is too strong, appropriate evaluation does not enter the mental workspace when it should, and one does not study. Once in the mental workspace, if the thought is deemed to be accurate ("Yes, you did not study enough"), the system changes its plan; namely, to study instead of going to see the movie.

Consider a case of anxiety, where the same thought arises: "You didn't study enough!" If gating is too weak, the thought arises too easily or comes in more frequently, moving it toward the spectrum of intrusive events. If there is an excessively strong gate, the thought does not enter consciousness and one does not feel a need to reevaluate the present plan, leading to the unhappy result that one does not study when indeed one should. Once in the mental workspace, if evaluation of the thought is inaccurate (you have indeed studied

Table 17.A1 Modes of failure.

	Normal	Pathological
Gating	Appropriate	Too weak, too strong
Salience	Appropriate	Too little, too much, etc.
Evaluation	Accurate	(In)accurate
Prediction error	Accurate	(In)accurate, failure to reset

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

enough when you think you didn't), the thought can become pervasive, lead to excessive worry, or inappropriately change the task to studying.

Salience signals can be faulty when studying seems much more appropriate than necessary, and the task is changed from going to a movie to studying. Another example of inappropriate salience occurs when one is studying, and studying appears so much more important than sleep that one continues studying until the moment of the exam, paradoxically hampering performance.

An inaccurate prediction error signal arises when one is on task, but nonetheless gets a signal that one is off task. For example, when someone is studying for a test, one feels anxiety despite engaging in studying, which can paradoxically undermine test preparation through perseverative worry (instead of studying).

An additional pathway that can compound mode failures emerges from reinforcement learning to combine with the faulty elements described above. If the threshold of the gate is too low, the interrupt changes the ongoing plan. The change itself (because it is new) can be reinforcing. This reinforcement leads to increased salience of the event that produced the change in plan. As salience iteratively increases, the likelihood increases that the interrupt that is linked to the new plan is going to disrupt other plans in the future (positive feedback loop). If the new plan is now in place, the stronger salience of the interrupt supports its maintenance. For example, with perseverative worry, the worry led to a change in plan (didn't go to the movie), which then led the person to study more and indeed feel better, which then functions as a reward signal that leads to a reinforcement of perseverative worry in the future. Such a simple mechanism could account for the common finding that OCD worsens with time. An important question for future research into effective interventions will be how to rein in this positive feedback loop afforded by reinforcement learning.

From "Intrusive Thinking: From Molecules to Free Will," edited by Peter W. Kalivas and Martin P. Paulus. Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1

From “Intrusive Thinking: From Molecules to Free Will,” edited by Peter W. Kalivas and Martin P. Paulus.
Strüngmann Forum Reports, vol. 30, Julia R. Lupp, series editor. Cambridge, MA: MIT Press. ISBN 978-0-262-54237-1