Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Behavioral Sciences

From tripping and falling to ruminating and worrying: a meta-control account of repetitive negative thinking

Peter F Hitchcock[1,2] and Michael J Frank[2,3,*]

Repetitive negative thinking (RNT) is a transdiagnostic construct that encompasses rumination and worry, yet what precisely is shared between rumination and worry is unclear. To clarify this, we develop a meta-control account of RNT. Meta-control refers to the reinforcement and control of mental behavior via similar computations as reinforce and control motor behavior. We propose rumination and worry are coarse terms for failure in meta-control, just as tripping and falling are coarse terms for failure in motor control. We delineate four meta-control stages and risk factors increasing the chance of failure at each, including open-ended thoughts (stage 1), individual differences influencing subgoal execution (stage 2) and switching (stage 3), and challenges inherent to learning adaptive mental behavior (stage 4). Distinguishing these stages therefore elucidates diverse processes that lead to the same behavior of excessive RNT. Our account also subsumes prominent clinical accounts of RNT into a computational cognitive neuroscience framework.

Addresses
[1] Department of Psychology, Emory University, Atlanta, GA, United States
[2] Cognitive, Linguistic, and Psychological Sciences, Brown University, Providence, RI, United States
[3] Carney Institute for Brain Science, Brown University, Providence, RI, United States

Corresponding author: Hitchcock, Peter F (peter.hitchcock@emory.edu)
[*] Twitter account: @PF_Hitchcock, @LnccBrown

Current Opinion in Behavioral Sciences 2024, **0**:103156

This review comes from a themed issue on **Cognitive Flexibility**

Edited by: **Roshan Cools** and **Lucina Uddin**

Available online xxxx

Received: 1 November 2023; Revised: 4 January 2024; Accepted: 22 January 2024

https://doi.org/10.1016/j.cobeha.2024.101356

2352–1546/© 2024 Elsevier Ltd. All rights reserved.

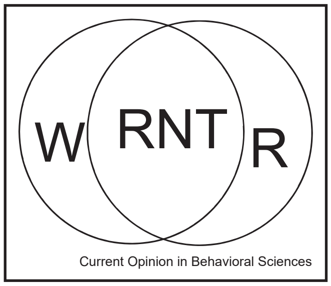
Introduction

Why do we dwell on our flaws and past mistakes and fixate on the uncertainty of the future? And what do these tendencies have in common? An emerging consensus in clinical science is that perseverating on the past (rumination) and fixating on the uncertain future (worry) are intimately related — in fact, that both thinking patterns can be (partly) subsumed by a higher-order construct, such as repetitive negative thinking (RNT) [1–4]. This consensus is relatively new: for decades, rumination and worry had been thought of as both conceptually distinct and as differential risk factors for depressive [5] and generalized anxiety [6] disorders, respectively (reviewed in Ref. [4•]). Yet, RNT is now understood as a transdiagnostic process: it is a risk factor for, and elevated within, most Axis I disorders [2,7], although most related to depression and anxiety [8]. And rumination and worry are now recognized as more similar than distinct [1,2,4,9]: factor analysis reveals that many rumination and worry items load onto a common factor that is more strongly associated with depression and anxiety than specific rumination and worry factors [9–13]; but see also Ref. [14••]. The definitions of rumination and worry also clearly overlap: both thinking patterns are repetitive, difficult to control, and negatively valenced (although with somewhat different content, such as focus on the past vs future; e.g. [1,2,4]; Figure 1).

In short, there is mounting psychometric, conceptual, and diagnostic evidence that rumination and worry are related. Yet, what is missing is an interdisciplinary basic scientific framework that makes testable predictions about the specific computations shared between rumination and worry. Here, we develop such a framework grounded in the computational decision sciences [16]. From a decision-making perspective, rumination and worry both reflect attempts to facilitate adaptive future decisions. Indeed, a rational agent would have no reason to perseverate on the past or prospect about the future unless they estimated that doing so would improve their current or future decisions [17]. RNT may be what happens when this typically adaptive process goes awry.

Our account, which we describe in detail below, offers some key advantages over prior work. First, it outlines conditions under which meta-control is more likely to fail (resulting in worry or rumination) versus when it is more likely to be productive (leading to adaptive problem solving or productive reflection). Hence, it overcomes the problem that some clinical accounts merely *stipulate* that rumination and worry are unproductive without clarifying why or how (see Refs. [3,18–20] for other clarifications). Second, by grounding RNT in

Figure 1



RNT refers to what is shared between worry (W) and rumination (R). This paper presents a computational account of what these thinking patterns share and also touches on how they diverge (see Refs. [4,8–11,13–15]).

computational cognitive neuroscience, our framework should help to organize findings on the neural circuits and neuromodulators implicated in RNT [21–23]. Finally, our framework identifies distinct pathways — ranging from neurocomputational individual differences to life stressors to learning challenges — that increase the chance of meta-control failure and thus excessive RNT. Crucially, excess RNT, especially by vulnerable individuals, appears to be a key pathway to episodes of depression, anxiety, or other psychopathologies at specific time points [24,25]. Insight into the distinct pathways that lead to excess RNT may therefore facilitate powerful and tailored interventions that can prevent these episodes or mitigate their damage.

Our account is specifically based on **meta-control** research from computational cognitive neuroscience [21–23]. Theoretical and empirical research and neural-network modeling in the last two decades have suggested that adaptive control over cognitive actions, such as gating content into and out of working memory, scaffold on top of reinforcement learning mechanisms that evolved to learn the consequences of, and guide the selection of, motor actions (e.g. [26–30]).

Here, we extend this insight to consider gating into working memory *hypotheses* with the potential to initiate rumination or worry, such as "I am socially incompetent" or "My spouse is ill." Based on prior work [26,31], we identify **gating in** an overarching task as the first of four stages of meta-control (Figure 2). We take 'rumination' and 'worry' to be generic labels for what happens when one or more stages of meta-control fail, leading to an unproductive episode of thinking (just as 'tripping' and 'falling' are generic labels for what happens when one or more stages of motor control fail, leading to errant motor behavior[1]). We equate rumination and worry with failure because these thinking patterns are often considered unproductive by definition [5,6] (although we acknowledge that other accounts use these terms differently; e.g. [3,20,32]).

Four stages of meta-control failures leading to rumination or worry

We identify four meta-control stages [26,31] at which failure may occur. Figure 2 introduces the four stages with the straightforward example of mental multiplication (left side) followed by hypothesis evaluation that might lead to RNT (right side). The remainder of this paper reviews the four stages and risk factors that make failure more likely at each stage. Our account highlights that numerous prominent clinical theories of RNT focus on a single stage of meta-control and forges links to decision science at each stage. It is thus unifying and consilient.

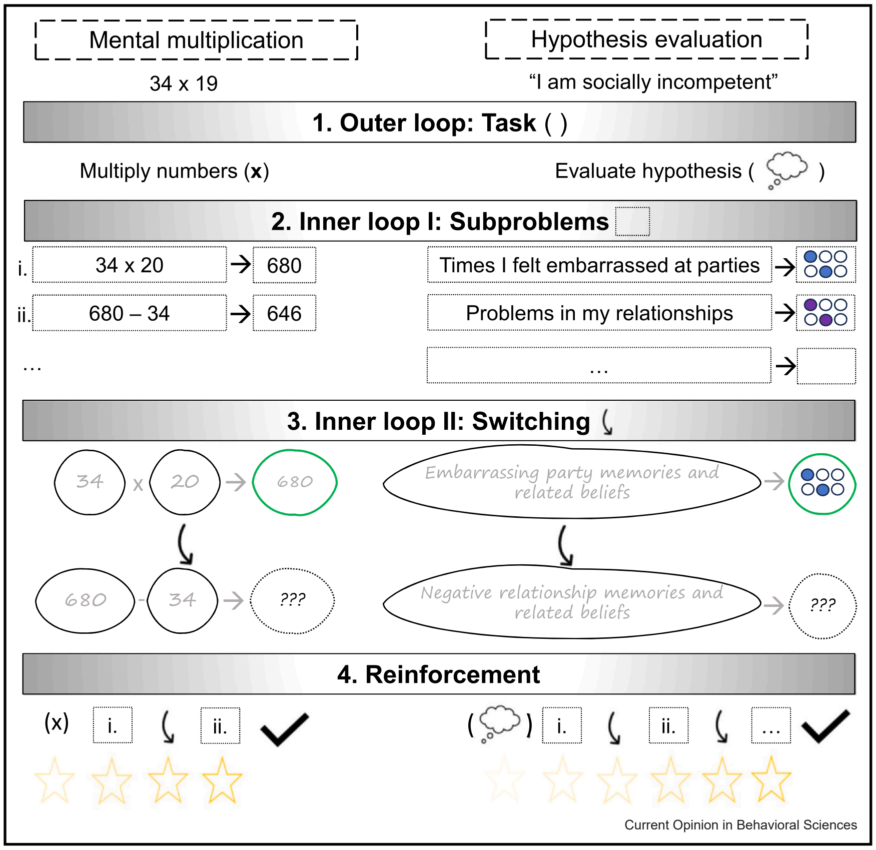
Stage 1. of meta-control is to gate and maintain an overarching task goal in mind. In the left scenario, you must remember that you are multiplying 34 and 19 (rather than subtracting, adding, or mentally rearranging them). In the right (affective, self-relevant) scenario, you are trying to evaluate the hypothesis "I am socially incompetent."

Stage 2. of meta-control is to complete a sequence of subproblems in the service of the overarching task. In the left scenario, completing just two subproblems solves the task: (1) finding the easy product 34×20 and (2) subtracting 19 from it. In the right scenario, in contrast, the set of subproblems is much more open ended: you might (1) recall some memories and beliefs about other times you felt embarrassed at parties, (2) think of problems that you have had in your relationships, and (...) focus on any number of other emotionally or topically pertinent subproblems — each of which might elicit further subproblems. (In the figure, the circles represent mental activity patterns constituting different memories and beliefs represented by different colors.)

Stage 3. of meta-control is to switch between subproblems. Switching requires sufficient confidence that the previous subproblem has been solved or addressed. In the left scenario, you must be confident that 34 times 20 is 680 to move on to subtracting 34 from that product. In the right scenario, you must be sufficiently satisfied with the memories and beliefs that come to mind related to embarrassing yourself at parties to move on to relationship problems.

[1] One way that this analogy is imperfect is that tripping and falling typically refer to rather extreme failures of motor control, whereas rumination and worry refer to a wide range of unproductive thinking (including not just extended spells of unproductive negative thinking but also a few moments 'spiraling' about a potential negative event). The key analogy we wish to make is therefore simply between errant motor behavior and errant mental behavior; we use tripping and falling simply as convenient examples.

Figure 2



The four stages of meta-control illustrated with mental multiplication (left) and extended to affective hypothesis evaluation that might instigate rumination (right). This figure illustrates the stages of meta-control in two scenarios. In the left scenario, you are challenged to solve a difficult mental multiplication problem in your head (34 × 19). In the right scenario, you are at a party conversing with someone who suddenly interrupts you, claiming that they need to use the restroom — leading you to consider the hypothesis "I am socially incompetent."

Stage 4. of meta-control is reinforcement of the operations that constituted the episode. Sometimes, reinforcement might be external. In the left scenario, the person who challenged you to mentally multiply might say, "You've got it! 646 is the answer." In the right scenario, you might see evidence that seems to confirm your hypothesis, such as that your former conversation party has begun a new conversation without having stopped at the restroom. Yet, reinforcement can also presumably come internally — such as if you are confident that you have solved the multiplication problem, as well as from beliefs, internalized norms or goals (see Ref. [33]), and replays of an experience that seem to support your hypothesis [34]. Whatever its source, it may be more difficult to ascribe reinforcement to stages of meta-control far in time from its receipt, especially when many mental operations occur before receipt. (In the figure, reinforcement is denoted with a checkmark and decreasing credit ascribed to earlier operations via fading gold stars.)

Stage 1 (outer loop): hypothesis selection

This stage refers to the overarching task that guides an episode of meta-control. In meta-control research (e.g. [26,27]), the overarching task (outer loop) is often experimentally defined, and sometimes it is also straightforward in everyday life — such as when mentally multiplying two numbers (Figure 2, right; although see Box 1 for discussion of complications for even these problems). Yet, many overarching tasks, such as evaluating a hypothesis about oneself (Figure 2, right), are more nebulous and open ended.[2] An influential clinical account argues that RNT is especially likely when "general, superordinate, and decontextualized mental

[2] Of note, our framework does not require that overarching hypotheses (e.g. "I am socially incompetent") are always consciously accessible. For instance, cognitive theory suggests that automatic thoughts that often come to mind emanate from core beliefs that may be difficult to identify. Indeed, part of the work of cognitive therapy is to learn to recognize (and thus make more consciously accessible) these core beliefs (e.g. [35,36]).

Box 1 The challenge of defining and reinforcing subproblems

The main text focuses on the distinct challenges of completing subproblems related to concrete versus open-ended hypotheses, treating the set of subproblems as predefined. For instance, it assumes that the two subproblems shown in (Figure 2, left) will be the ones pursued to solve the mental multiplication problem. In reality, defining subproblems is a challenging problem in its own right even for concrete hypotheses to say nothing of more open-ended ones. For instance, one could attempt to solve the multiplication problem the long-form way or via another multistep strategy, such as using three steps all involving products and sums: $34 \times 10 + 30 \times 9 + 4 \times 9$.

That multiple subproblems can be pursued in the service of the same goal poses further challenges for credit assignment (stage 4). For instance, variable strategies will lead to varying completion times of the overarching task; different subproblems will themselves vary in completion time due to fatigue and practice; and different strategies may lead to entirely different outcomes for the same overarching task. For instance, an attempt to solve 34×19 via a long-form strategy might lead to continued forgetting and needing to restart and ultimately giving up — even for someone capable of solving the problem with a smarter strategy, such as the two-step solution in Figure 2.

In the case of affective open-ended hypotheses, the challenge of defining a set of subproblems compound, making credit assignment yet more challenging. A variety of other complications also come into play: for instance, negatively valenced hypotheses might trigger hard-wired **Pavlovian responses** that prevent gating and thus preclude subgoal execution [42] or undermine instrumental reinforcement (if one believes that an overarching negative hypothesis has been confirmed — and the resulting conclusion feels awful). Such interactions could be investigated experimentally in paradigms where the confirmation of a hypothesis can be dissociated from the valence of the outcome (e.g. [43]).

Subproblem selection and its consequences for reinforcement are also relevant to intervention. Repeatedly practicing mental behaviors learned in therapy (e.g. cognitive defusion, asking 'how' rather than 'why' questions, Socratic questioning) may make them easier to implement, and thereby decrease variability in delays and increase the chance of consistent success — leading to more consistent reinforcement. Indeed, computational modeling clarifies that meta-learning an effective control or gating policy can enhance the efficiency of computations and thereby maximize reward. Specifically, the policies that are easiest to learn via credit assignment may be those that are more efficient to use because they involve fewer and clearer subgoals to complete [26,27,44,45]. One implication for psychotherapy theory is that certain skills might be recast as subproblems of others, rather than as wholly distinct strategies. For instance, cognitive defusion and asking how versus why questions might be initial skills that pave the way for more complex strategies, such as Socratic questioning. This provides an alternate framing for debates about how distinct certain strategies are from each other [39] and raises the possibility of shaping complex skills by first establishing more basic ones [45].

representations" ([37], p. 260) guide (sub)goal or action selection [3,18•,37]. From a meta-control perspective, a key issue with open-ended hypotheses is that they lead to numerous subproblems (see Figure 2, right, stage 2) — none of which have much chance of definitively resolving the hypothesis itself or informing specific external actions. In contrast, concrete hypotheses (such as "I angered my friend at the party last night") lead to more well-defined subproblems and associated actions (such as "I wonder if my friend did not like it when I teased her partner?" and plans to change specific future behavior).

Although selecting concrete (vs open-ended) hypotheses may often lead to more tractable subproblems, concrete hypotheses can of course elicit more open-ended ones ("I angered my friend at the party" might lead to thinking "I am socially incompetent"). Notably, **postevent processing** [34], which has been formalized in terms of upward counterfactuals [38•], is implicated in social anxiety and other psychopathology — implying that the tendency to revisit specific event details is a risk factor in its own right. Nonetheless, our account highlights that such thinking is most likely to be prolonged and unproductive when thoughts stray from specific situations and events and become open ended.

Open-ended hypotheses pose a couple of specific challenges from a computational perspective. For one, the inferences involved in such hypotheses (such as about one's own and others' traits, about dynamics like gossip propagating through complex social networks) may be intractably complex [39]. Second, when a hypothesis spawns many subproblems, this presumably makes it difficult to robustly maintain the hypothesis itself in mind [26] and to ascribe consequences to it [40••] (see Box 1 and stage 4).

The fraught nature of open-ended hypothesizing is well recognized. Indeed, key strategies in RNT treatments are to promote concrete thinking (thinking grounded in specific events and details and focused on 'how' instead of 'why' questions) and to track the deleterious consequences of open-ended thinking (e.g. [41]). Our account lays basic scientific groundwork for such strategies by identifying hypothesis evaluation as the first stage of meta-control. This should help to elucidate how open-ended hypothesizing interacts with risk factors at other stages.

Stage 2 (inner loop 1): executing subgoals

Once subgoals have been defined, they must be executed — and certain individual differences in execution may predispose toward RNT. Some differences may comparably increase the tendency to ruminate or worry. For instance, **trait neuroticism** may predispose toward RNT simply because more negative memories and polarized beliefs come to mind, irrespective of whether they concern the past or the future. Interestingly, self-reported 'stickiness' of thinking predicts task disengagement [46] and may be a subjective marker of protracted subgoal execution. Other individual differences may predispose specifically to rumination or worry (see Ref. [14]). For instance, self-referential thinking is more core to the construct of rumination than worry and thus may have a stronger relationship to it (**divergent validity**). Recent years have seen advances in techniques for investigating self-judgment, ranging from computational modeling to event-related neuroimaging to real-world prediction (e.g. [47–50]). Yet, most work focused on depression rather than rumination (but see Ref. [51]) and no work of which we are aware examined **divergent validity**. Conversely, uncertainty aversion is more core to the construct of worry than rumination. Thus, individual differences underlying this tendency — in the tendency to differentiate within semantic space [52•], for instance — may relate more to worry than rumination, although again divergent-validity studies are lacking.

Stage 3 (inner loop 2): switching between subgoals

Once a subgoal is completed, it must be recognized as such so that the next subgoal can begin. Recognition of subgoal completion is relatively straightforward in mental multiplication but less so for open-ended hypotheses, which may lead subproblems to be processed in a disorganized way and with much revisiting (Figure 2).

Influential clinical theories posit that rumination and worry arise from a difficulty inhibiting, or switching between, mental representations. Stage 3 is one stage where this may pose difficulties — in completing subgoals and switching to new ones (see also Ref. [53•]). Yet, there is disagreement about whether such difficulties occur with all types of content or only negatively valenced content — in which case it is questionable whether the problem is switching *per se* [54–57]; see also Ref. [25,58]. Whitmer and Gotlib [55] reviewed evidence suggesting that rumination may relate, in part, to a trait-like proclivity for stability over flexibility. The idea is that individual differences in this balance are neither uniformly good nor bad but rather are helpful in different settings: stability is helpful when it is necessary to concentrate on a focal task and avoid distractions, and flexibility is helpful when task switching is needed. Yet Whitmer and Gotlib [55] noted that higher trait-like stability will increase the propensity to ruminate when it is coupled with individual differences (e.g. neuroticism) or contexts (e.g. distressing life events) that increase negative affect (especially given that negative affect itself may stabilize cognition [55]).

It is still an open question whether RNT relates to valence-independent differences in stability versus flexibility in a way that can be consistently assessed behaviorally — especially given considerable practical challenges with such assessment (e.g. [59]) evident since the literature review in the study by Whitmer and Gotlib [55] and that most research has focused on rumination rather than worry or transdiagnostic RNT. One reason this question matters is that dopamine appears to regulate stability versus flexibility balance, with higher striatal dopamine linked to flexibility and higher prefrontal dopamine linked to stability [60]. Such individual differences may inform tailored RNT treatments.

One way to stop perseverating on open-ended hypotheses besides completing subgoals is to engage with something external, such as an attention-demanding task [61] or a cue to potential reward [62]. Yet, when the propensity for RNT becomes entrenched, it may interfere with external tasks [46,63], including learning about reinforcement contingencies [64,65•] (reviewed in Refs. [18,62,66]). Disrupted learning is an especially insidious effect because it not only alters behavior in the moment but also in related situations in the future and because it may set off a feedback loop of withdrawal from the external world and preoccupation by RNT ([65•] for discussion).

Stage 4: learning from consequences

Crucially, meta-control concerns not only the control of mental behavior but also *learning its consequences* through reinforcement so that (ideally) mental behavior becomes more adaptive through learning [21,26,28]. Thus, the final stage[3] of meta-control is learning the consequences of the mental operations performed. Many clinical accounts of RNT view it as governed by the principles of operant conditioning [67–70]; hence, meta-control is a natural formal framework [40••].

One key learning challenge is that unlike motor actions, mental behaviors entail carrying out one or more mental operations. Operations are thus necessarily interposed between the initiation of a mental action and the receipt of its outcome (for instance, subproblems are interposed between gating a hypothesis into mind and receiving reinforcement about the gating operation itself). In a task that directly compared the ability to learn optimal mental versus motor behaviors through trial and error, we found that people had more difficulty learning optimal mental behaviors. Thus, it may be especially difficult to learn the consequences of the operations that constitute RNT — and to learn adaptive thinking patterns more generally [40••]. The intrinsic challenge of learning about mental behavior might help explain why people are apt to neglect its negative consequences, as posited by many RNT theories [67–70]. Whereas Hitchcock and Frank [40••] considered learning about simple mental operations that were well defined in our task (to isolate the challenge of learning adaptive cognitive actions even in a simple case), as noted computational modeling has shown that credit

[3] Of note, this is simply the final stage within our conceptual framework; in the actual process of meta-control, learning may take place at various points (i.e. it does not necessarily happen only once the overarching task attempt has been completed).

assignment is more challenging when multiple subproblems are required (stage 1) and multiple strategies can be used, some of which are more readily ascribed credit than others (Box 1) [26,27,44,45]. There also may be individual differences in learning from mental behaviors per se; we are currently investigating this possibility.

These learning challenges may be one reason that people are so prone to acquiring mental behavior with poor long-run consequences (such as **gating in** abstract hypotheses that might instigate RNT). An additional challenge, besides difficulty learning RNT's negative consequences, is that RNT may actually have some positive consequences — which might outweigh its negative consequences (especially if they are more difficult to learn). Indeed, numerous sources of reinforcement for RNT (i.e. **secondary gains**) have been proposed, such as that people ruminate to conform to unjust gender norms [5,71], because it facilitates social bonding (e.g. through co-rumination [72]), or that it justifies abstention from their commitments [68]. One prominent account argues that RNT functions to pre-emptively lower affect so that when negative experiences occur, they do not lead to a precipitous drop in affect [73]. This theory could be formalized via reinforcement learning if RNT decreases affect by lowering expectations, so that subsequent prediction errors are more positive. Increasing the plausibility that RNT functions to deflate expectations in this way, prediction errors — which are defined as outcomes minus expectations — have a strong influence on momentary mood [74]. However, the mechanics of ascribing reinforcement to an action that is aversive, yet which is believed to preclude a more aversive outcome, are unclear (but see, e.g. [75,76], for general evidence that counterfactual outcomes and not choosing an action both influence reinforcement learning).

Interactions between stages

A key advantage of our framework is that it allows theorizing about how individual differences will affect multiple stages and how stages interact. As one example, here we focus on how similar individual differences may influence both overarching hypotheses (stage 1) and subgoal execution (stage 2). Above, we discussed how negative self-judgment can influence subgoal execution (stage 2), yet a heightened tendency toward self-referential thinking is also tantamount to a tendency to consider open-ended hypotheses (stage 1) — given that self-schemas are by definition abstractions devoid of specific context [77]. From a computational perspective, decontextualized representations are thought to arise from the gradual incorporation of representations of specific, context-laden experiences stored in the hippocampal system into decontextualized representations of what is shared between them in neocortex [78]. One reason to build such representations is to enable recognizing that superficially distinct situations share contingencies (e.g. stimulus-response) [27]; see also Ref. [79••] for discussion. Although this may often be adaptive, overgeneralization is also possible, such as when long-held negative views about oneself lead to an inaccurate assumption of poor efficacy in new situations [77].

Past computational research on self-judgment has focused on judgments and memories separately from learning (e.g. [47–50]). And research on the inference of shared learning contingencies has used arbitrary stimuli (e.g. fractal images) when studying situations that may share a (hidden) contingency structure (e.g. [27,80]). One novel direction would be to join these lines of research and examine how specific individual differences in self-referential processing (first examined separately) influence the inference of shared contingencies when self-referential (instead of arbitrary) stimuli are used as cues to shared structure.

Life events may also influence both stages. For instance, Martin and Tesser [20] construed rumination as fundamentally tied to goal progress: when a specific goal is blocked (such as failing to find a part while assembling a desk), the resulting repetitive thinking will be very specific, whereas thinking will be more abstract and open ended in the face of setbacks (e.g. job loss) that influence many goals at once (stage 1). Different overarching problems will also lead to different subgoals (stage 2). For instance, life events involving death and separation are more predictive of depression than generalized anxiety symptoms, whereas those involving loss and danger are more predictive of generalized anxiety than depression [81]. This makes it plausible that distinct event types also differently predict rumination and worry (see Refs. [24,25]). Our framework clarifies that although the content of the problems elicited by different life events will diverge, major life stressors will have a shared effect at the process level, namely, promoting open-ended thinking.

Conclusion

We developed a basic scientific framework for understanding rumination and worry as failed meta-control. Of note, by introducing our framework in this brief article format, we were only able to cite rather than review some of the key literature that informed it (e.g. the neural bases of meta-control and RNT, executive function differences that might confer vulnerability for and/or result from RNT, methods of measuring RNT). The scope of our framework is also currently limited to RNT alone; we did not attempt to locate it within the wider constructs of perseverative or repetitive thought (e.g. [4]) or intrusive thinking and low mindfulness (e.g. [15,82]). Also regarding scope: our framework focuses on

RNT as active attempts to resolve a hypothesis gone awry, rather than passive RNT arising from habitual ineffective thinking patterns (although it is naturally compatible with habit-focused accounts, e.g. [67], and reinforcement learning formalizations thereof).

Notwithstanding these limitations, we outlined a framework that establishes consilience between research on RNT and the decision sciences; elucidates diverse risk factors that can each lead to excess RNT; and lays a basic science groundwork for specialized treatments [83] for rumination [41], worry [84], and transdiagnostic RNT [85••,86]. Our framework makes testable predictions about how distinct processes (e.g. individual and/or situation-based differences in proneness to open-ended thinking, trait-like stability over flexibility, challenges of learning about mental behavior especially when compounded by open-ended thinking) may lead to the same behavioral output of excessive RNT.

Funding
P.F.H. was supported by National Institute of Mental Health (United States) Grant F32MH123055. M.J.F. was supported by National Institutes of Mental Health (United States) Grants P50MH119467 and R01MH084840-08A.

Data Availability

No data were used for the research described in the article.

Declaration of Competing Interest

The authors declare no conflicts of interest with regard to the present manuscript.

Acknowledgements
Thank you to Rachel Bedder and Paul Sharp for many productive discussions that contributed to the ideas in this paper and to Rachel Bedder and Marta Migó for feedback on the manuscript.

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest.

1. Harvey AG, Watkins E, Mansell W: **Cognitive Behavioural Processes Across Psychological Disorders: A Transdiagnostic Approach to Research and Treatment**. Oxford University Press; 2004.

2. Ehring T, Watkins ER: **Repetitive negative thinking as a transdiagnostic process**. *Int J Cogn Ther* 2008, **1**:192-205.

3. Watkins ER: **Constructive and unconstructive repetitive thought**. *Psychol Bull* 2008, **134**:163-206.

4. • Hallion LS, Wright AGC, Joormann J, Kusmierski SN, Coutanche MN, Caulfield MK: **A five-factor model of perseverative thought**. *J Psychopathol Clin Sci* 2022, **131**:235-252.

Dimensional analysis of clustering among online participants' ratings of recent preoccupying thoughts, which found support for five factors underlying perseverative thought: dyscontrol, interpersonal, self-focus, valence, and uncertainty. Perseverative thought here is similar to repetitive negative thought as discussed in the main text, but with a wider scope because it also includes positively valenced thoughts.

5. Nolen-Hoeksema S: **Sex differences in unipolar depression: evidence and theory**. *Psychol Bull* 1987, **101**:259-282.

6. Borkovec TD, Robinson E, Pruzinsky T, DePree JA: **Preliminary exploration of worry: some characteristics and processes**. *Behav Res Ther* 1983, **21**:9-16.

7. Wahl K, Ehring T, Kley H, Lieb R, Meyer A, Kordon A, *et al.*: **Is repetitive negative thinking a transdiagnostic process? A comparison of key processes of RNT in depression, generalized anxiety disorder, obsessive-compulsive disorder, and community controls**. *J Behav Ther Exp Psychiatry* 2019, **64**:45-53.

8. Aldao A, Nolen-Hoeksema S, Schweizer S: **Emotion-regulation strategies across psychopathology: a meta-analytic review**. *Clin Psychol Rev* 2010, **30**:217-237.

9. Funk J, Takano K, Schumm H, Ehring T: **The bi-factor model of repetitive negative thinking: common vs. unique factors as predictors of depression and anxiety**. *J Behav Ther Exp Psychiatry* 2022, **77**:101781.

10. Ehring T, Zetsche U, Weidacker K, Wahl K, Schönfeld S, Ehlers A: **The Perseverative Thinking Questionnaire (PTQ): validation of a content-independent measure of repetitive negative thinking**. *J Behav Ther Exp Psychiatry* 2011, **42**:225-232.

11. McEvoy PM, Mahoney AEJ, Moulds ML: **Are worry, rumination, and post-event processing one and the same? Development of the repetitive thinking questionnaire**. *J Anxiety Disord* 2010, **24**:509-519.

12. McEvoy PM, Salmon K, Hyett MP, Jose PE, Gutenbrunner C, Bryson K, *et al.*: **Repetitive negative thinking as a transdiagnostic predictor of depression and anxiety symptoms in adolescents**. *Assessment* 2019, **26**:324-335.

13. Spinhoven P, van Hemert AM, Penninx BW: **Repetitive negative thinking as a predictor of depression and anxiety: a longitudinal cohort study**. *J Affect Disord* 2018, **241**:216-225.

14. •• Stade EC, Ruscio AM: **A meta-analysis of the relationship between worry and rumination**. *Clin Psychol Sci* 2023, **11**:552-573.

Meta-analysis of rumination and worry that made some substantial improvements to prior meta-analyses (for instance, by examining previously unreported rumination and worry correlations and examining subtypes of rumination). It found that, although rumination and worry were highly correlated, they were not entirely overlapping. This finding is consistent with our view in the main text—that rumination and worry overlap substantially but also diverge in important ways (e.g., see Figure 1 and the in-text discussion of the need to establish divergent validity for processes that may relate more to rumination vs. worry and vice versa).

15. Gustavson DE, du Pont A, Whisman MA, Miyake A: **Evidence for transdiagnostic repetitive negative thinking and its association with rumination, worry, and depression and anxiety symptoms: a commonality analysis**. *Collabra: Psychol* 2018, **4**:1-18.

16. Frömer R, Nassar MR : **Belief Updates, Learning and Adaptive Decision Making**; 2023. doi:〈10.31234/osf.io/qndba〉.

17. Gershman SJ, Horvitz EJ, Tenenbaum JB: **Computational rationality: a converging paradigm for intelligence in brains, minds, and machines**. *Science* 2015, **349**:273-278.

18. • Watkins ER, Roberts H: **Reflecting on rumination: consequences, causes, mechanisms and treatment of rumination**. *Behav Res Ther* 2020, **127**:103573.

Comprehensive recent review of rumination that covers its effects, theorized cognitive components, and developments in treatment. It proposes a conceptual model that ties together many of the cognitive components.

19. Segerstrom SC, Stanton AL, Alden LE, Shortridge BE: **A multidimensional structure for repetitive thought: what's on**

your mind, and how, and how much? *J Pers Soc Psychol* 2003, **85**:909-921.

20. Martin LL, Tesser A: **Some ruminative thoughts**. *Ruminative Thoughts* 1996, **9**:1-47.

21. Dayan P: **How to set the switches on this thing**. *Curr Opin Neurobiol* 2012, **22**:1068-1074.

22. Eppinger B, Goschke T, Musslick S: **Meta-control: from psychology to computational neuroscience**. *Cogn Affect Behav Neurosci* 2021, **21**:447-452.

23. Westbrook A, Braver TS: **Dopamine does double duty in motivating cognitive effort**. *Neuron* 2016, **91**:708.

24. Nolen-Hoeksema S, Watkins ER: **A heuristic for developing transdiagnostic models of psychopathology: explaining multifinality and divergent trajectories**. *Perspect Psychol Sci* 2011, **6**:589-609.

25. Hitchcock PF, Fried EI, Frank MJ: **Computational psychiatry needs time and context**. *Annu Rev Psychol* 2022, **73**:243-270.

26. O'Reilly RC, Frank MJ: **Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia**. *Neural Comput* 2006, **18**:283-328.

27. Collins AGE, Frank MJ: **Cognitive control over learning: creating, clustering, and generalizing task-set structure**. *Psychol Rev* 2013, **120**:190-229.

28. Braver TS, Cohen JD: **On the control of control: the role of dopamine in regulating prefrontal function and working memory**. In *Attention and Performance XVIII*. Edited by Monsell S, Driver J. MIT Press; 2000:713-738.

29. • Trutti AC, Verschooren S, Forstmann BU, Boag RJ: **Understanding subprocesses of working memory through the lens of model-based cognitive neuroscience**. *Curr Opin Behav Sci* 2021, **38**:57-65.

An updated review of literature suggesting that working-memory operations are governed by computations similar to those involved in learning motor actions.

30. Rac-Lubashevsky R, Frank MJ: **Analogous computations in working memory input, output and motor gating: electrophysiological and computational modeling evidence**. *PLoS Comput Biol* 2021, **17**:e1008971.

31. Todd M, Niv Y, Cohen JD: **Learning to use working memory in partially observable environments through dopaminergic reinforcement**. *Adv Neural Inf Process Syst* 2008, **20**:1700-1707.

32. Andrews PW, Thomson JA Jr.: **The bright side of being blue: depression as an adaptation for analyzing complex problems**. *Psychol Rev* 2009, **116**:620-654.

33. Molinaro G, Collins AGE: **A goal-centric outlook on learning**. *Trends Cogn Sci* 2023, **27**:1150-1164, https://doi.org/10.1016/j.tics.2023.08.011

34. Brozovich F, Heimberg RG: **An analysis of post-event processing in social anxiety disorder**. *Clin Psychol Rev* 2008, **28**:891-903.

35. Beck JS: **Cognitive behavior therapy**. Basics and Beyond. edn 3, Guilford Publications; 2020.

36. Berwian I, Hitchcock P, Pisupati S, Schoen G, Niv Y: **Using learning theories to advance psychotherapy theory and research**. *PsyArXiv* 2023,1-20, https://doi.org/10.31234/osf.io/vbp6c

37. Watkins E: **Dysregulation in level of goal and action identification across psychological disorders**. *Clin Psychol Rev* 2011, **31**:260-278.

38. • Hunter LE, Meer EA, Gillan CM, Hsu M, Daw ND: **Increased and biased deliberation in social anxiety**. *Nat Hum Behav* 2022, **6**:146-154.

In a social reinforcement learning game, social-anxiety symptoms predicted responsivity to upward counterfactuals—that is, to what could have gone better.

39. FeldmanHall O, Nassar MR: **The computational challenge of social learning**. *Trends Cogn Sci* 2021, **25**:1045-1057.

40. •• Hitchcock PF, Frank MJ: **The challenge of learning adaptive mental behavior**. *PsyArXiv* 2023,1-69, https://doi.org/10.31234/osf.io/agprs.

Direct comparison of the ability to learn optimal mental vs. motor behaviors (e.g., add/subtract numbers vs. select letters based on the spatial position top/bottom, respectively). Overall, people were worse at learning optimal mental (vs. motor) behaviors, which computational modeling traced to impaired credit assignment for mental actions. Despite the overall impairment at the group level, a subset of participants actually learned optimal mental behaviors better than motor behaviors. It is currently unclear whether this heterogeneity is meaningful—for instance, whether the challenge of performing mental behaviors actually enhances learning for a subset of people, mirroring other paradoxical effects in which difficulty enhances reinforcement learning..

41. Watkins ER: **Rumination-Focused Cognitive-Behavioral Therapy for Depression**. Guilford Publications; 2018.

42. Dayan P, Huys QJM: **Serotonin, inhibition, and negative mood**. *PLoS Comput Biol* 2008, **4**:e4.

43. van den Bosch R, Lambregts B, Määttä J, Hofmans L, Papadopetraki D, Westbrook A, *et al.*: **Striatal dopamine dissociates methylphenidate effects on value-based versus surprise-based reversal learning**. *Nat Commun* 2022, **13**:4962.

44. Frank MJ, Badre D: **Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis**. *Cereb Cortex* 2012, **22**:509-526.

45. Krueger KA, Dayan P: **Flexible shaping: how learning in small steps helps**. *Cognition* 2009, **110**:380-394.

46. Huijser S, Verkaik M, van Vugt MK, Taatgen NA: **Captivated by thought: "sticky" thinking leaves traces of perceptual decoupling in task-evoked pupil size**. *PLoS One* 2020, **15**:e0243532.

47. Hitchcock PF, Britton WB, Mehta KP, Frank MJ: **Self-judgment dissected: a computational modeling analysis of self-referential processing and its relationship to trait mindfulness facets and depression symptoms**. *Cogn Affect Behav Neurosci* 2023, **23**:171-189.

48. Allison GO, Kamath RA, Carrillo V, Alqueza KL, Pagliaccio D, Slavich GM, *et al.*: **Self-referential processing in remitted depression: an event-related potential study**. *Biol Psychiatry Glob Open Sci* 2023, **3**:119-129.

49. Dainer-Best J, Lee HY, Shumake JD, Yeager DS, Beevers CG: **Determining optimal parameters of the self-referent encoding task: a large-scale examination of self-referent cognition and depression**. *Psychol Assess* 2018, **30**:1527-1540.

50. Cataldo AM, Scheuer L, Maksimovskiy AL, Germine LT, Dillon DG: **Abnormal evidence accumulation underlies the positive memory deficit in depression**. *J Exp Psychol Gen* 2023, **152**:139-156.

51. Auerbach RP, Stanton CH, Proudfit GH, Pizzagalli DA: **Self-referential processing in depressed adolescents: a high-density event-related potential study**. *J Abnorm Psychol* 2015, **124**:233-245.

52. • Vives M-L, de Bruin D, van Baar JM, FeldmanHall O, Bhandari A: **Uncertainty aversion predicts the neural expansion of semantic representations**. *Nat Hum Behav* 2023, **7**:765-775.

Shows that people who report being averse to uncertainty show evidence of less similarity between semantic representations in various tasks and an expanded neural space during semantic processing.

53. • Sharp PB, Russek EM, Huys QJM, Dolan RJ, Eldar E: **Correction: humans perseverate on punishment avoidance goals in**

multigoal reinforcement learning. *Elife* 2022, **11**:1-25, https://doi.org/10.7554/eLife.83998.

A clever task design wherein participants attempted to maximize reward/minimize punishment based on instructions that changed on a trial-wise basis about which of two features (colored circles) would lead to reward or punishment. Participants accessed the features by selecting one of two fractal images that probabilistically led to the presence of each feature, where {image-feature} probabilities drifted slowly and were independent from each other. The key findings were, first, that participants used a less flexible/computationally simpler strategy to avoid punishment than reward, which led to the persistence of punishment avoidance into reward-pursuit trials and blocks; and, second, in exploratory findings, trait worry was associated with particular difficulty disengaging from the goal of avoiding punishment.

54. Gotlib IH, Joormann J: **Cognition and depression: current status and future directions**. *Annu Rev Clin Psychol* 2010, **6**:285-312.

55. Whitmer AJ, Gotlib IH: **An attentional scope model of rumination**. *Psychol Bull* 2013, **139**:1036-1061.

56. Berenbaum H: **An initiation–termination two-phase model of worrying**. *Clin Psychol Rev* 2010, **30**:962-975.

57. Koster EHW, De Lissnyder E, Derakshan N, De Raedt R: **Understanding depressive rumination from a cognitive science perspective: the impaired disengagement hypothesis**. *Clin Psychol Rev* 2011, **31**:138-145.

58. Grahek I, Shenhav A, Musslick S, Krebs RM, Koster EHW: **Motivation and cognitive control in depression**. *Neurosci Biobehav Rev* 2019, **102**:371-381.

59. Enkavi AZ, Eisenberg IW, Bissett PG, Mazza GL, MacKinnon DP, Marsch LA, *et al*.: **Large-scale analysis of test–retest reliabilities of self-regulation measures**. *Proc Natl Acad Sci* 2019, **116**:5472-5477.

60. Cools R, D'Esposito M: **Inverted-U–shaped dopamine actions on human working memory and cognitive control**. *Biol Psychiatry* 2011, **69**:e113-e125.

61. Hertel PT, Rude SS: **Depressive deficits in memory: focusing attention improves subsequent recall**. *J Exp Psychol Gen* 1991, **120**:301-309.

62. Rutherford AV, McDougle SD, Joormann J: **"Don't [ruminate], be happy": a cognitive perspective linking depression and anhedonia**. *Clin Psychol Rev* 2023, **101**:102255.

63. Jones NP, Siegle GJ, Muelly ER, Haggerty A, Ghinassi F: **Poor performance on cognitive tasks in depression: doing too much or not enough?** *Cogn Affect Behav Neurosci* 2010, **10**:129-140.

64. Whitmer AJ, Frank MJ, Gotlib IH: **Sensitivity to reward and punishment in major depressive disorder: effects of rumination and of single versus multiple experiences**. *Cogn Emot* 2012, **26**:1475-1485.

65. • Hitchcock P, Forman E, Rothstein N, Zhang F, Kounios J, Niv Y, *et al*.: **Rumination derails reinforcement learning with possible implications for ineffective behavior**. *Clin Psychol Sci* 2022, **10**:714-733.

Dysphoric individuals showed impaired learning, in an attention-demanding reinforcement-learning task, when learning was punctuated by cues to ruminate vs. to engage in neutral visual imagery (in a within-subject design). Interestingly, there was a greater reduction in learning in the rumination (vs. neutral) condition among less dysphoric individuals—whereas high-dysphoric individuals showed poor learning irrespective of condition. Another key finding was that trait rumination correlated with a computational-model parameter interpretable as a narrower scope of attention during reinforcement learning. This is consistent with findings that rumination corresponds to narrower (and less flexible but more stable) attention (ref. [55] and see discussion in Stage 3, main text) and is the first such finding (of which we are aware) in a learning context.

66. Reilly EE, Lavender JM, Berner LA, Brown TA, Wierenga CE, Kaye WH: **Could repetitive negative thinking interfere with corrective learning? The example of anorexia nervosa**. *Int J Eat Disord* 2018, **52**:36-41, https://doi.org/10.1002/eat.22997

67. Watkins ER, Nolen-Hoeksema S: **A habit-goal framework of depressive rumination**. *J Abnorm Psychol* 2014, **123**:24-34.

68. Nolen-Hoeksema S, Wisco BE, Lyubomirsky S: **Rethinking rumination**. *Perspect Psychol Sci* 2008, **3**:400-424.

69. Behar E, DiMarco ID, Hekler EB, Mohlman J, Staples AM: **Current theoretical models of generalized anxiety disorder (GAD): conceptual review and treatment implications**. *J Anxiety Disord* 2009, **23**:1011-1023.

70. Newman MG, Llera SJ: **A novel theory of experiential avoidance in generalized anxiety disorder: a review and synthesis of research supporting a contrast avoidance model of worry**. *Clin Psychol Rev* 2011, **31**:371-382.

71. Nolen-Hoeksema S: **Emotion regulation and psychopathology: the role of gender**. *Annu Rev Clin Psychol* 2012, **8**:161-187.

72. Rose AJ: **The costs and benefits of co-rumination**. *Child Dev Perspect* 2021, **15**:176-181.

73. Newman MG, Rackoff GN, Zhu Y, Kim H: **A transdiagnostic evaluation of contrast avoidance across generalized anxiety disorder, major depressive disorder, and social anxiety disorder**. *J Anxiety Disord* 2023, **93**:102662.

74. Rutledge RB, Skandali N, Dayan P, Dolan RJ: **A computational and neural model of momentary subjective well-being**. *Proc Natl Acad Sci USA* 2014, **111**:12252-12257.

75. Palminteri S, Khamassi M, Joffily M, Coricelli G: **Contextual modulation of value signals in reward and punishment learning**. *Nat Commun* 2015, **6**:8096.

76. Ben-Artzi I, Kessler Y, Nicenboim B, Shahar N: **Computational mechanisms underlying latent value updating of unchosen actions**. *Sci Adv* 2023, **9**:eadi2704.

77. Disner SG, Beevers CG, Haigh EAP, Beck AT: **Neural mechanisms of the cognitive model of depression**. *Nat Rev Neurosci* 2011, **12**:467-477.

78. McClelland JL, McNaughton BL, O'Reilly RC: **Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory**. *Psychol Rev* 1995, **102**:419-457.

79. •• Bedder RL, Pisupati S, Niv Y: **Modelling rumination as a state-inference process**. *Proc Annu Meet Cogn Sci Soc* 2023, **45**:1-7.

A reinforcement-learning account of rumination that casts it as an attempt to resolve the current state of the environment (among known possible states) and finds conditions under which repeated sampling to do so is optimal. Although the paper focuses on inference between established states, it touches on the need to construct such states in the first place—a process that we link in the main text to the formation of self-schemas (which will also be developed in a forthcoming paper by Bedder,* Hitchcock,* and Sharp). More generally, the state-centered focus in this line of work is complementary to our focus here on mental action learning, selection, and control; both link RNT to rich areas of reinforcement learning (e.g., meta-control, representation learning, POMDPs).

80. Collins AGE, Frank MJ: **Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning**. *Cognition* 2016, **152**:160-169.

81. Kendler KS, Hettema JM, Butera F, Gardner CO, Prescott CA: **Life event dimensions of loss, humiliation, entrapment, and danger in the prediction of onsets of major depression and generalized anxiety**. *Arch Gen Psychiatry* 2003, **60**:789-796.

82. Kalivas PW, Paulus MP: **Intrusive Thinking: From Molecules to Free Will**. MIT Press; 2021.

83. Hitchcock PF: **Of strong swords and fine scalpels: developing robust clinical principles to cut through heterogeneity**. *J Psychopathol Clin Sci* 2023, https://doi.org/10.31234/osf.io/tvk4n In Press.

84. Orsillo SM, Roemer L: **Acceptance- and Mindfulness-Based Approaches to Anxiety: Conceptualization and Treatment**. Springer Science & Business Media; 2007.

85. •• Joubert AE, Grierson AB, Li I, Sharrock MJ, Moulds ML, Werner-Seidler A, *et al*.: **Managing rumination and worry: a randomised controlled trial of an internet intervention targeting repetitive negative thinking delivered with and without clinician guidance**. *Behav Res Ther* 2023, **168**:104378.
Found that short (3-session) self- and clinician-guided versions of a transdiagnostic RNT intervention (The Managing Rumination and Worry Program) were more effective at reducing RNT, rumination, and worry than treatment as usual at post-treatment and a 3-month follow-up. The clinician-guided treatment (which was also an online treatment but included clinical guidance through phone support) was more effective than the self-guided one. Following ref. [41], the intervention translated basic clinical science findings related to RNT (reviewed in ref. [18]) and associated techniques into a short online transdiagnostic intervention “in the form of an illustrated comic-style story that follows two fictional characters who learn to better manage rumination and worry” (p. 4)—thereby creating an intervention with high potential for scalability.

86. Ruiz FJ, Peña-Vargas A, Ramírez ES, Suárez-Falcón JC, García-Martín MB, García-Beltrán DM, *et al*.: **Efficacy of a two-session repetitive negative thinking-focused acceptance and commitment therapy (ACT) protocol for depression and generalized anxiety disorder: a randomized waitlist control trial**. *Psychotherapy* 2020, **57**:444-456.

87. Apšvalka D, Ferreira CS, Schmitz TW, Rowe JB, Anderson MC: **Dynamic targeting enables domain-general inhibitory control over action and thought by the prefrontal cortex**. *Nat Commun* 2022, **13**:274.

Glossary

Divergent validity: Theoretically meaningful discriminability of one measure from another.

Gating in: According to computational neuroscience models and experimental data (e.g. [26,28–31]), people can ‘gate’ content into working memory, allowing it to be stored in a more robust form in the face of distracting or task-irrelevant content. Gating policies can be acquired via reinforcement learning (e.g. if entertaining and maintaining a joke in mind leads to actually telling the joke and people laughing in response, these processes will be more likely to be repeated in the future). Gating is assumed to be determined by prior experience and the current context and thus may happen partly out of voluntary control, although it also may be modulated by partly voluntary processes, such as thought suppression (e.g. [87]).

Meta-control: Learning via reinforcement, and controlling the execution of, policies for mental behavior (e.g. working memory operations).

Pavlovian responses: ‘Hard-wired’ approach or avoid responses that can sometimes conflict with instrumental ones (such as withdrawing from a negatively valenced stimulus even if approaching it would lead to reward, or failing to entertain a negative hypothesis even if it would be in one’s best interest to do so).

Postevent processing: Threatening appraisal of a social or other event after it has occurred.

Secondary gains: Positive consequences of behaviors that are maladaptive on balance.

Trait neuroticism: The stable tendency to experience strong and frequent negative reactions.