Search for solutions, learning, simulation, and choice processes in suicidal behavior

Alexandre Y. Dombrovski | Michael N. Hallquist

Abstract
Suicide may be viewed as an unfortunate outcome of failures in decision processes. Such failures occur when the demands of a crisis exceed a person's capacity to (i) search for options, (ii) learn and simulate possible futures, and (iii) make advantageous value-based choices. Can individual-level decision deficits and biases drive the progression of the suicidal crisis? Our overview of the evidence on this question is informed by clinical theory and grounded in reinforcement learning and behavioral economics. Cohort and case–control studies provide strong evidence that limited cognitive capacity and particularly impaired cognitive control are associated with suicidal behavior, imposing cognitive constraints on decision-making. We conceptualize suicidal ideation as an element of impoverished consideration sets resulting from a search for solutions under cognitive constraints and mood-congruent Pavlovian influences, a view supported by mostly indirect evidence. More compelling is the evidence of impaired learning in people with a history of suicidal behavior. We speculate that an inability to simulate alternative futures using one's model of the world may undermine alternative solutions in a suicidal crisis. The hypothesis supported by the strongest evidence is that the selection of suicide over alternatives is facilitated by a choice process undermined by randomness. Case–control studies using gambling tasks, armed bandits, and delay discounting support this claim. Future experimental studies will need to uncover real-time dynamics of choice processes in suicidal people. In summary, the decision process framework sheds light on neurocognitive mechanisms that facilitate the progression of the suicidal crisis.
1 | INTRODUCTION

This article discusses how features of decision processes may predispose people to suicidal behavior. We relate the decision process account to the traditional clinical view in which a combination of stable traits and transient states increases the risk of suicide. We explore how cognitive capacity limitations preclude people from finding and properly considering alternative solutions in a crisis, or from fully considering the implications of suicide itself. The paper is written for a diverse audience interested in behavior and decision-making, including clinical psychologists, psychiatrists, and cognitive scientists. While our perspective is grounded in formal models of decision-making, particularly reinforcement learning (RL), we assume only a basic knowledge of probability and statistics. We provide nonmathematical intuitions for key concepts and refer the interested reader to recent reviews (Gureckis & Love, 2015; Hallquist & Dombrowski, 2020; Patzelt et al., 2018) and Sutton and Barto's authoritative book (2018). Similarly, readers unfamiliar with suicide research can reference Box 1 for definitions, Table 1 for risk factors (further reading suggestions in the legend), and Table 2 for methodological considerations; in the text, we reference reviews of existing theories of suicide and of empirical studies of risk factors.

As with any complex behavior, nonoverlapping and noncompeting levels of explanation have been invoked to explain suicide, from genetic and epigenetic (Cheung et al., 2020; Sokolowski & Wasserman, 2020) to social (Biddle et al., 2008; Calati et al., 2019; Gould, 2001), anthropological (Lawrence et al., 2016), and environmental (Beskow et al., 1994; Yip et al., 2012). Depression, addiction, and psychosis precipitate suicidal behavior whereas maladaptive personality traits are thought of as predisposing factors. Research on trait predispositions provides a broad account of emotional experiences and behavioral tendencies that facilitate suicidal behavior. However, to understand relevant cognitive processes and their neural substrates we need a formal, mechanistic account. Our paper begins with a historical introduction, discusses the current view of the psychological diathesis of suicide, and finally presents the decision process framework, focusing on (i) how people construct a set of options in a crisis, (ii) how they simulate possible outcomes and update the value of suicide versus alternatives, and (iii) the process of choosing the best solution. This framework is a collection of ideas intended to inform empirical research rather than one particular theory of suicide. This overview does not cover neural underpinnings of suicidal behavior. When speaking of dissociable cognitive processes, however, we discuss neural circuits that subserve them, hoping to inform neurobiological studies of suicidal behavior.

BOX 1 Concepts in suicide research

Terminology

Suicide: The act of killing oneself deliberately initiated and performed by the person concerned in the full knowledge or expectation of its fatal outcome (World Health Organization, 1998).

Suicide attempt: A potentially self-injurious behavior with a nonfatal outcome, for which there is evidence (either explicit or implicit) that the person intended at some (nonzero) level to kill himself/herself. A suicide attempt may or may not result in injuries (O'Carroll et al., 1996).

Suicidal behavior: Here, suicide or suicide attempt.

Suicidal ideation: Any self-reported thoughts of engaging in suicide-related behavior (O'Carroll et al., 1996).

Attempt lethality: Physical damage and danger to life resulting from a suicide attempt (Beck et al., 1975).

Can suicide be rational? Today, many embrace Cicero's view of suicide as rational (Section 2.1), and assisted suicide has been legalized, with safeguards, in Benelux, Switzerland, Columbia, Canada, and a few US and Australian states. The public emphasis on precipitating factors -- economic hardship, social adversity, or physical illness -- may lead one to think that most suicides are rational. Yet, overall, the data speak to the contrary: psychological autopsy studies find that 90% of suicide victims (vs. 27% of controls) suffer from mental disorders (J. T. O. Cavanagh et al., 2003). Furthermore, people's preference for suicide over alternatives changes dynamically in ways that they themselves often cannot predict (Section 4).
### TABLE 1  Some established risk factors for suicide: stress-diathesis perspective

| Predisposing (distal) | Precipitating (proximal) |
|-----------------------|--------------------------|
| Male sex (WHO Mental Health, 2016) | Psychopathology: depression, psychosis, addiction |
| Family history of suicidal behavior (meta-analysis: Franklin et al., 2017) | Social stressors: conflict, loss of relationship, bereavement, unemployment, financial and legal problems (Agbayewa et al., 1998; Agerbo et al., 2002; Coope et al., 2015; Duberstein et al., 2004; Luoma & Pearson, 2002; Milner et al., 2017) |
| Personality traits (Section 3) | Access to firearms and lethal drugs (Beautrais et al., 1996; Birckmayer & Hemenway, 2001; Conwell et al., 2002; Milner et al., 2017) |
| Childhood trauma (meta-analysis: Zatti et al., 2017) | Social contagion in young people (systematic review: Niedzwiedz et al., 2014) |
| Cognitive deficits (Section 4) | Physical illness and pain (Juurlink et al., 2004; Waern et al., 2002) |
| Societal factors: (lack of) religious injunctions, cultural attitudes toward suicide (Colucci & Martin, 2008; Lawrence et al., 2016) | |

*Note: Several recent papers review the risk factors for suicide (Bentley et al., 2016; Franklin et al., 2017; Hawton et al., 2013; Turecki et al., 2019). The stress-diathesis model distinguishes distal or predisposing from proximal or precipitating factors (Mann, 2003; Turecki et al., 2019, p. 201).*

### TABLE 2  Predicting a rare outcome: design approaches

| Approach | Strengths | Limitations and ways to address them |
|----------|-----------|--------------------------------------|
| Cohort studies | Yield true risk estimates. Less prone to confounds. | Given the low rate of suicide (1:10,000 per year; WHO Mental Health, 2016), adequately powered in-depth cohort studies of death by suicide are prohibitively expensive and require follow-up durations on the order of decades. |
| Longitudinal studies of samples enriched for suicide risk | True risk estimates, but higher power than cohort studies. | Suicide rates in enriched samples still rarely exceed 1% per year (Ballard et al., 2016; Wenzel et al., 2011), and the ethical imperative to prevent suicides whenever possible, reduces this low base rate even further. Enrichment for past suicidal behavior results in survivors of low-lethality attempts being over-represented. Since past suicide attempts predict future suicidal behavior, it is hard to know whether to control for them statistically and focus on partial effects of substantively interesting predictors or credit these predictors for shared variance. |
| Retrospective case–control studies | Adequate power. Shorter duration. | Cannot provide true risk estimates. Prone to confounds and require careful consideration of psychopathology (e.g., nonsuicidal patient comparison groups) and other confounds (e.g., medication exposure and brain injury from suicide attempts). |
| Surrogate outcomes, from passive death wish and suicidal ideation with or without a plan to attempted suicide<sup>a</sup> | Higher base rate and power | Not necessarily representative of death by suicide. Suicide attempts can be rendered more representative of death by suicide by sampling or separating statistically near-fatal (high-lethality) suicide attempts. seriousness of suicide attempts varies greatly with age, with attempt completion ratio falling from >100 in younger adults to the single digits in older men (CDC, 2020). Thus, attempted suicide in old age provides a window into factors that lead to suicide. |

<sup>a</sup>Terms implying a purpose and especially individual responsibility in suicidal behavior—committing suicide, completed suicide, failed attempt, suicide gesture—are often eschewed as pejorative. We agree with the need to be sensitive and generally follow this principle, although established technical terms such as “attempt: completion ratio” are difficult to replace. We also feel that if an objectionable perspective is sufficiently common among researchers or practitioners, it deserves to be explicitly refuted rather than censored. Finally, one cannot understand suicide without delving into the dark side of the human nature, creating the need to acknowledge, non-judgmentally and without exception, all human behaviors, experiences and motivations.
2 | BACKGROUND

2.1 | Early ideas

Ancient thinkers were mostly concerned with philosophical arguments for and against the moral permissibility of suicide. One other theme that emerges in Plato (Laws, IX) and certainly in the Roman Stoics is that of weighing pros and cons of living versus dying. Cicero says it plainly: “When a man’s circumstances contain a preponderance of things in accordance with nature, it is appropriate for him to remain alive; when he possesses or sees in prospect a majority of the contrary things, it is appropriate for him to depart from life” (Cicero, III, 60–61, cited in Cholbi, 2013). Another idea—that suicide may be caused not by tragic circumstances, but by rash judgment—is conveyed in the tragic story of Pyramus and Thysbe in Ovid’s Metamorphoses, the source for the plot of Romeo and Juliet.

The notion that suicide results from personality traits and the manner in which decisions are made is elaborated in Dostoevsky’s novels. In Crime and Punishment, the aging landowner Svidrigailov pursues life’s pleasures at any material and moral cost, a path that ends in premeditated suicide when the pleasures are run through. In some sense, Svidrigailov’s suicide extends a lifelong pattern of behavior. A complex character, one could certainly think of Svidrigailov’s depravity, cheating at cards, cynicism, manipulativeness, lack of remorse and recklessness as features of personality disorders, or externalizing psychopathology (cf. moral insanity of the 19th century). We will return to the contributions of externalizing to suicide in Section 3, which discusses the personality trait perspective.

By contrast, the decision perspective is revealed to us by the omniscient narrator through the thoughts and experiences of the protagonist, Raskolnikov, a St. Petersburg law school dropout. Raskolnikov resolves to break out of desperate poverty by murdering his pawnbroker, a murder later revealed to be a metaphor for suicide. Raskolnikov ends up murdering not only the repulsive old woman—aiming to end his deprivation and humiliation—but also, accidentally, her pregnant sister embodying innocence and hope. He later explains to his lover: “Was it the old crone that I killed? I killed myself, not the old crone!” (Dostoevsky, 2012). Before resolving his moral doubts about the murder, Raskolnikov learns that the pawnbroker will be alone one night and decides to act. He struggles with details, such as how to obtain and dispose of the axe. Dostoevsky remarks: “And, indeed, if it had ever happened that everything to the least point could have been considered and finally settled, and no uncertainty of any kind had remained, he would, it seems, have renounced it all as something absurd, monstrous and impossible. But a whole mass of unsettled points and uncertainties remained” (part 1, chapter 6, Dostoevsky, 2012). The uncertainty, the “mass of unsettled points” paralyze Raskolnikov’s judgment and obscure otherwise insurmountable deterrents. Strikingly, the murder is set into motion by a circumstance irrelevant to the fundamental judgment about its moral worth and likely consequences, or its long-term value in the parlance of decision science. (Raskolnikov kills himself in the original 1865 draft of the novel; Svidrigailov is one of Raskolnikov’s alter egos embodying the dark side of human nature. A similar pair, the predator Stavrogin and the suicide victim Kirillov, appears in Dostoevsky’s Demons.) We will return to the role of uncertainty and constraints on decision-making in the suicidal crisis in Section 4, which presents the decision process framework.

2.2 | Modern clinical theory: insights and limitations

Modern psychological accounts of suicide such as the entrapment (Williams et al., 2005) and escape (Baumeister, 1990) frameworks, interpersonal theory (Van Orden et al., 2010), motivational-volitional theory (O’Connor & Kirtley, 2018), and the three-step theory (Klonsky et al., 2018) share the view of suicide as a flight from intolerable affective states (for a useful, brief review see O’Connor & Nock, 2014). All of these theories seek to explain the suicidal process, the progression from emotional distress to contemplating suicide, and finally to suicidal action. The putative factor that facilitates the transition to action is termed capability for suicide. These theories provide a rich phenomenological framework for thinking about suicide and are rooted in the same 19th-century descriptions of pre-suicidal states that Dostoevsky likely drew on (Efremov, 2008). Millner et al. (2020) have recently highlighted some of their limitations, including their modest predictive power and the vagueness of their predictions. The latter relates to the crud factor (Meehl, 1990), a term referring to the shared variance between new psychological constructs (e.g., hopelessness) and previously proposed constructs (depression). In fact, shared variance among constructs of entrapment, defeat, hopelessness, psychic pain, perceived burdensomeness, thwarted belongingness, and depressive symptoms reaches 50% (Cramer et al., 2019; DeLisle & Holden, 2009), making it hard to demonstrate that these specific states rather than cognitive symptoms of depression lead to thoughts of suicide. More importantly, descriptive theories view the progression from depression to suicidal
behavior both as a key outcome and a unique dynamic psychological process. As a result, most of their constructs are specific to suicide or depression, and external explanatory frameworks are not invoked. The circularity of this approach precludes us from looking for lower-order mechanisms underlying the suicidal process that can be related to broader accounts of behavior. In summary, existing accounts of suicide are phenomenological rather than mechanistic, and new perspectives grounded in basic, generally applicable models of human behavior are needed. Here, we propose that formal learning theory and decision neuroscience provide a productive framework for understanding cognitive and decision processes involved in suicide.

3 | THE ROLE OF PREDISPOSITIONS AND PREFERENCES IN SUICIDE

As noted in Box 1, the research literature on suicide is replete with associational studies of risk factors ranging from military service to marital discord (Steele et al., 2018). Although such studies can inform broadband assessments of risk in clinical settings, discerning causal risk factors from proxies can be hard (cf. Kraemer et al., 2001). Evidence of relatively stable individual predispositions to suicide, however, may provide useful insight into mechanisms of risk. In particular, we focus on the role of impulsivity and neuroticism, personality traits that have stable neurobiological correlates and are rooted in early temperament (Lahey, 2009; Silverman et al., 2019).

Neuroticism is the tendency to experience frequent and intense negative emotions such as anger, sadness, or anxiety. In the psychopathology literature, such symptoms have been called “internalizing” and relate to depression, anxiety disorders, eating disorders (Griffith et al., 2010). Impulsivity is the propensity to engage in careless, reward-dependent, and short-sighted behavior. Impulsivity is an important component of the broader externalizing spectrum, which includes symptoms such as substance abuse, aggression, and antagonistic behaviors (Krueger et al., 2002). Meta-structural models of psychopathology note that most symptoms of psychopathology can be related to internalizing and externalizing spectra, suggesting that these may be shared liability factors for mental illness (Kotov et al., 2017).

Neuroticism and impulsivity have been consistently associated with suicidal behavior. For example, in a large military sample, Nock et al. (2018) found that difficulty controlling suicidal thoughts and “tempting fate” (more extreme risk-taking), which likely relate to neuroticism and impulsivity, respectively, were associated with a greater probability of attempt after controlling for several other risk factors. Likewise, in a large population sample, Peters et al. (2018) found that neuroticism was prospectively associated with completed suicide over a 10-year follow-up period. More broadly, the associations of depression, hopelessness, and neuroticism with suicide can be understood in terms of the internalizing spectrum, which emerges in meta-analysis as an important liability factor (Franklin et al., 2017), though its effect may be more robust in early-onset suicidal behavior (Szücs et al., 2020). Finally, there is preliminary evidence that the interaction between internalizing and externalizing psychopathology facilitates suicidal behavior (Duprey et al., 2020), consistent with the motivational-volitional theory of suicide (O’Connor & Kirtley, 2018) and internalizing–externalizing comorbidity in borderline personality disorder (BPD; Eaton et al., 2012).

Illustrated by the character of Svidrigailov above, externalizing symptoms including antisocial personality features and substance abuse predict suicide attempts above and beyond neuroticism/internalizing (Franklin et al., 2017; Verona et al., 2004). A broader interpretation of the motivational-volitional and three-step theories of suicide (Section 2.2) is that internalizing primarily promotes the emergence of suicidal ideas, and externalizing, suicidal action. On the other hand, one would predict that if impulsivity played a causal role in suicide, suicidal acts among highly impulsive people would be marked by a lack of premeditation. Paradoxically, this is not the case (see Anestis et al., 2014), suggesting that the apparent association between impulsivity and suicide may be better explained by shared externalizing liability. For example, externalizing, but not internalizing, features in BPD prospectively predict life stressors that are at least partly generated by the individual (e.g., relationship break-up; T. A. Allen, Dombrowski, Soloff, & Hallquist, 2020), suggesting that externalizing may also lead to suicide via the accumulation of negative life events that can develop into a suicidal crisis.

Despite the clinical utility of internalizing and externalizing in suicide risk assessment, studies of these features have relied almost exclusively on self-report methods such as dimensional surveys or categorical diagnoses including depression and alcohol use disorder. The challenge is that internalizing and externalizing are broad, heterogeneous constructs and when assessed by survey measures, they provide general information distal from the factors that propel a suicidal crisis (cf. Anestis et al., 2014; Sharma et al., 2014). Herein lies a tension between an idiosyncratic focus on the unique circumstances of a given crisis and the scientific motivation to identify nomothetic risk factors that are associated with suicide across people. Stable individual differences in personality constructs may serve as a backdrop for the development
of a crisis, but the choice of suicide over alternatives is often made within hours or minutes of an attempt (Millner et al., 2017). Thus, a framework for elaborating the interface between personality factors (and other predispositions) and the suicidal process must specify how predispositions play out in the moment to alter decision-making and propel or avert a crisis.

Below, we argue that decision theory and RL provide such a framework. Where do individual differences in personality fit into this view? DeYoung (2015) notes that personality traits are, “probabilistic descriptions of relatively stable patterns of emotion, motivation, cognitive, and behavior, in response to classes of stimuli” (p. 35). This definition underscores that traits can influence the tendency to experience a state like anger in a given situation (i.e., traits influence momentary experience probabilistically). It also highlights that traits shape individual responses to specific stimuli. For example, a person who is more extraverted may respond to social attention by becoming garrulous, while a person high in rejection sensitivity may perceive such attention as a threat and become aloof. Thus, knowing something about an individual’s personality may tell us something about how she/he will respond to specific circumstances that could be part of the progression to suicide.

Returning to neuroticism and impulsivity, we argue that studying the interface between decision process and individual differences in traits can illuminate how stable predispositions alter risk for suicide. That said, pinning down the behavioral manifestations of personality traits is not straightforward, particularly in the case of impulsivity. Studies of the associations between laboratory measures (e.g., Go-No Go tasks or the Balloon Analog Risk Task) self-report measures of impulsivity often find weak and/or inconsistent effects (Creswell et al., 2019; Sharma et al., 2014). We propose that a decision process framework allows one to test more specific hypotheses about how personality variables relate to behavior in the lab (for a more detailed account, see T. A. Allen, Schreiber, Hall, & Hallquist, 2020). As discussed in Section 4.4.1, we propose that neuroticism is linked with lower expected value of alternative actions to suicide—particularly those requiring more effort or multiple steps—during a crisis. Mood-dependent impulsivity, on the other hand, may promote ineffective attempts to escape intolerable emotional states, including suicidal behavior (Millner et al., 2019). Short-sighted intertemporal preferences have been suggested to facilitate impulsive suicidal behavior, a hypothesis discussed in Section 4.6.3.

4 | SUICIDE AS A COGNITIVE FAILURE

If suicide follows a decision, as people have thought since antiquity, how precisely is this decision made? We recall Dostoevsky’s description of Raskolnikov’s decision process, with its inconsistency and confusion. This section discusses decision deficits and biases that may favor suicide over alternatives in the crisis. The association of suicidal behavior with disadvantageous real-life decisions such as addiction and gambling (Vijayakumar et al., 2011; Wong et al., 2010) suggests a shared decisional impairment. One possible mechanism of this impairment consists of broad cognitive constraints on decision-making, and particularly deficits in cognitive control (the ability to prioritize task-relevant processing): we review the evidence supporting such constraints (Section 4.1). We also discuss studies directly demonstrating impaired decision-making in suicidal behavior using the Iowa Gambling Task (IGT) (Section 4.2). Finally, we consider decision process disruptions corresponding to theoretical stages of the suicidal process: emergence of active suicidal thoughts (consideration set construction), weighing of suicide versus alternative solutions (learning and simulation), and eventual choice (Sections 4.3–4.7).

4.1 | Evidence for the cognitive diathesis

We are interested in cognitive constraints because they potentially explain the inability of suicidal people to fully consider alternatives and consequences of suicide. The growing empirical literature linking suicidal behavior to cognitive deficits has been recently reviewed and subjected to meta-analyses (K. J. D. Allen et al., 2019; Bredemeier & Miller, 2015; Cha et al., 2019; Jollant et al., 2011; Liu et al., 2017; Richard-Devantoy, Berlim, & Jollant, 2014; Richard-Devantoy et al., 2015; Rutter et al., 2020; Saffer & Klonsky, 2018). Two general patterns can be discerned. First, the association of suicidal behavior with deficits in global intelligence and certain memory and cognitive control (executive) domains is robust, but the effect sizes are quantitatively modest. Second, the magnitude of deficits scales inversely with the strength of comparison groups (healthy subjects, nonsuicidal patients, suicide ideators with no history of attempts; see Table 2 for background on study designs; Richard-Devantoy, Berlim, et al., 2014; Richard-Devantoy et al., 2015;
Saffer & Klonsky, 2018). Most cognitive measures do not differentiate between people who only endorse suicidal ideas and those who act on them (Saffer & Klonsky, 2018), important exceptions being Stroop inhibition and decision-making tasks discussed in the later sections. That said, as discussed later (Section 4.4), cognitive deficits associated with suicidal ideation are of theoretical interest since they may constitute constraints on the construction of consideration sets.

The most general evidence of cognitive constraints comes from intelligence studies. At first, they appear mixed, but the negative association between intelligence and suicidal behavior is supported by large cohort studies, including national cohorts (Andersson et al., 2008; Gunnell et al., 2005, 2011) whereas contrary findings generally come from weaker cross-regional comparisons (Voracek, 2004, 2006a, 2006b) or smaller case–control studies (Park et al., 2015). Notably, lower intelligence is paradoxically protective against suicide in people with schizophrenia (Batty et al., 2010; Webb et al., 2011); our framework may not apply here.

The profile of deficits across traditional neuropsychological domains in suicide attempters is generally consistent with difficulties in generating and evaluating alternative solutions. Suicide attempters show impairments in long-term and autobiographical, but not short-term, memory (meta-analysis: Richard-Devantoy et al., 2015). It is possible that these deficits impair generation of alternative solutions in a crisis and memory-based simulations evaluating them, although evidence for this claim is modest (Evans et al., 1992; Pollock & Williams, 1998). Cognitive control or executive deficits are particularly toxic for learning and decision-making (Collins et al., 2017; Otto, Gershman, Markman, & Daw, 2013). Cognitive control impairments are seen in both suicide ideators and suicide attempters. Interestingly, Stroop inhibition is the only cognitive task to consistently distinguish suicide attempters from suicide ideators (Keilp et al., 2013; meta-analysis: Richard-Devantoy, Berlim, et al., 2014; Richard-Devantoy, Szanto, Butters, Kalkus, & Dombrovski, 2014). In summary, people who engage in suicidal behavior display deficits in global cognitive performance, memory and cognitive control above and beyond what is expected in psychopathology, but the stages of the suicidal process to which these deficits may contribute remain unclear. Also unclear is whether cognitive deficits merely contribute to a catastrophic accumulation of problems or propel the suicidal crisis once it begins. Finally, we assume that cognitive deficits operating during the suicidal crisis possess a trait component and are detectable at other times. Although not adequately tested in suicidal patients, this assumption appears reasonable based on studies of depression, which find deficits that persist in remission and are at best modestly moderated by current mood state (Bora et al., 2013; Rock et al., 2014; Snyder, 2013; however, see also McDermott & Ebmeier, 2009; Semkovska et al., 2019).

4.2 Initial studies of decision-making employing the IGT

After Jollant et al.’s (2005) report of poor IGT performance in suicide attempters, more than a dozen IGT studies followed, replicating the effect on the whole (meta-analysis: Richard-Devantoy, Berlim, et al., 2014). A recent, larger meta-analysis found that the effect persists after adding more studies (suicide attempters vs. healthy controls: Hedges $g = -0.54$, vs. patient controls: $g = -0.28$; Perrain et al., 2021). Notably, the effect size was smaller in recent compared to older studies, suggesting regression to the mean (Proteus phenomenon; Ioannidis & Trikalinos, 2005). While the IGT studies support the hypothesis of impaired decision-making in suicidal behavior, the nature of this impairment remains unclear. Successful performance on the IGT is defined as learning to prefer “good” decks of cards yielding small wins and even smaller losses over “bad” decks of cards bringing big wins and even bigger losses. The IGT is therefore potentially sensitive to deficits in both learning and value-based choice, leaving open the question of which of these functions is disrupted in suicidal behavior.

4.3 Overview of the decision process framework

In an effort to overcome some of the aforementioned limitations of existing theories of suicide, here, we consider the suicidal process from the perspective of decision theory and RL (Figure 1). People generally do not admit to considering suicide as an option when dealing with challenges of their lives; in the clinical jargon, they do not endorse suicidal ideation. However, even prior to any crisis, suicide and potential alternatives are assigned values based on experience and mental simulation of their consequences. Then, the person experiences a crisis, an unexpected event that brings about uncertainty and a threat to her core goals (Seeger et al., 1998). A crisis generates a sense of urgency to resolve the immediate problems through any means possible, often motivating vigorous action. Suicide may then be included in the consideration set (options among which one chooses in each instance) alongside other potential solutions, or actions,
based on their prior values and constraints: this is the emergence of suicidal ideation. The expected value (Box 2) of each action is updated through learning and simulation, as new information is obtained, outcomes are simulated and solutions are attempted. Choosing among actions based on their perceived values, the person may select inaction (appropriately or not), an ineffective action that perpetuates the crisis, or an action that resolves it. Below, we discuss each stage, focusing on individual characteristics or capacity limitations that may favor suicidal behavior over alternatives.

How could suicide get selected over alternatives in a crisis? Clinicians note that the decision to attempt suicide typically follows a limited consideration of the crisis, alternative solutions, and deterrents (Henriques et al., 2005; Shneidman et al., 1970). Ostensibly moderate triggers, such as an argument, can precipitate a suicidal act, overshadowing major deterrents, such as trauma to one’s family. While suicidal people often report internally debating reasons for living versus dying (Harris et al., 2010), most of those who go on to attempt suicide and survive are happy to be alive and regret their attempt (Henriques et al., 2005). These observations suggest that when choosing suicidal behavior, most people cannot accurately predict how the crisis will evolve, and their estimates of the relative value of suicide versus alternatives, including inaction, end up being inconsistent with their interests. There may be various causes for this inconsistency. Escape theories reviewed earlier emphasize intense negative affective states, which make a challenging situation feel catastrophic. As detailed in Section 4.4.1, Pavlovian influences may constitute one mechanism through which internal states color valuations (cf. Huys & Renz, 2017). The emergence of suicidal ideas can be understood in terms of selection of options into a consideration set based on precomputed values (Section 4.4.2). Another cause may be that suicidal people fail to integrate their moment-to-moment experiences with prior knowledge and values and to update their expectations flexibly (Sections 4.5.2 and 4.5.3). In the next section, we consider this possibility from the perspective of RL (Box 2; Sutton & Barto, 1998), one of the most successful formal accounts of human behavioral adaptation and maladaptation (for a general overview of clinical applications, see Hallquist & Dombrovski, 2020). We briefly introduce relevant concepts from RL (Section 4.5.1), discuss learning deficits that may be suicidogenic (Section 4.5.2), and review the evidence of learning deficits in suicidal behavior (Section 4.5.3). Unlike the constructs of psychological
theories, these processes, as well as choice processes discussed in the next section (Section 4.6), are mostly automatic and best accessed through behavioral and physiological experiments rather than interviews and self-report.

4.4 | Emergence of suicidal ideas. Pavlovian influences and consideration sets

The emergence of suicidal ideas seems to fall outside of a decision process framework. This conclusion may, however, be premature. From the decision theory standpoint, suicide is simply not part of the consideration set for most people, most of the time. Here, we first discuss Pavlovian processes that explain how negative affect may color prior values of possible solutions. Next, we discuss constraints on the search for solutions in a crisis.

4.4.1 | Mood-dependent valuation as Pavlovian influence or biased priors

Clinical theorists have long associated depression with the tendency to view the future with pessimism and to remember the past in overgeneralized and negative terms (Beck, 1963). From a decision theory standpoint, how would...
depressed mood and, more broadly, internalizing psychopathology shape the expected values of suicide vs. alternatives? We see three possibilities, which are not mutually exclusive. First, negative mood may reduce the probability of choosing an adaptive coping strategy via mood-dependent Pavlovian valuation (cf. Dayan et al., 2006). The Pavlovian system learns stimulus-outcome contingencies and is both insensitive to action outcomes and exquisitely sensitive to internal states. For example, the long-run value of hiring an attorney to file for bankruptcy may be trumped by the Pavlovian avoidance of the intense shame and self-reprisal associated with reviewing one’s overdrawn bank account. Relatedly, if suicide represents an escape from intolerable internal states, aversive Pavlovian cues (e.g., phone calls from debt collectors) may inflate the value of suicide as a Pavlovian-congruent response. In line with this hypothesis, Millner et al. (2019) found that suicidal individuals learned to respond more quickly to escape an aversive sound compared to nonsuicidal psychiatric patients, suggesting an escape bias that may reflect Pavlovian influences.

Second, negative mood (and internalizing more generally) may influence decision-making by setting pessimistic prior beliefs about the expected value of actions whose results are uncertain. Negative emotional states tend to promote inaction (likely also via Pavlovian influence; de Berker et al., 2016) and may generate a biased prior belief that any effortful attempt to cope with the crisis is hopeless (Huys et al., 2015). Why would beliefs about multiple options be altered simultaneously instead of being ruled out sequentially? Although standard RL models tend to focus solely on learning from experience about a single action or cue, there is increasing evidence that learning signals for a given action generalize to others and that this generalization is a governing process in the unfolding of mood (Eldar et al., 2016, 2018). Eldar’s model originally aimed to explain mood fluctuations, and has been extended to depression (Sharp et al., 2020). Altogether, we believe that Pavlovian valuation biases underlie internalizing psychopathology, depressing the expected values of alternatives to suicide and prior beliefs about the success of one’s own actions.

Third, emotional states likely favor suicidal ideas and undermine deterrents by altering what simulated future outcomes are considered. As noted by Huys and Renz (2017), emotions help to shape not only which actions are attempted, but also what sequences of events are simulated and evaluated, a meta-reasoning strategy that evokes the “tunnel vision” metaphor. It is thus likely that mood-incongruent versions of the future are overlooked both during the selection of possible solutions (Section 4.4.2) and their evaluation in the later stages the suicidal crisis (Section 4.5).

### 4.4.2 Cognitive constraints on consideration set construction

The construction of consideration sets—subsets of a larger universe of options entered into a final choice process—is an increasingly well understood predecisional cognitive process (Hauser, 2014; Hauser & Wernerfelt, 1990). When purchasing durable goods, for example, considering more options is more likely to result in a better choice up to a certain asymptote (Figure 2(a)). At the same time, the search for options takes time and effort. Thus the decision-maker should stop searching upon reaching the maximum net benefit (Figure 2, left, dashed line). When constructing a set of possible solutions under the new circumstances and uncertainty of a crisis, it would seem adaptive to select them in a flexible, situation-specific manner. Such flexibility, however, incurs additional search costs. Under cognitive constraints people reduce these costs by sampling options based on their cached (precomputed; Box 2) rather than situation-specific values (Phillips et al., 2019). For example, people tend to eliminate actions of low general, and particularly moral, value regardless of the situation, when asked to quickly name events that are possible, a tendency diminished when they are given more time to reflect (Phillips & Cushman, 2017). As noted above, Pavlovian influences generally depress values of mood-incongruent actions, making their selection less likely. Other factors influencing the general, rather than situation-specific, value of suicide such as religious injunctions or social models can operate at the early stage, suppressing its selection into the consideration set. This indeed seems to be the case for religious prohibitions, which reduce the prevalence of suicidal ideas, but do not necessarily prevent people from acting on them (Colucci & Martin, 2008).

The suicidal crisis is often marked by a sense of time pressure, triggered by imminent threats such as eviction or unmet financial obligations. Since the number of items selected into the consideration set is limited by time and cognitive resources fewer alternatives will be included when one is under time pressure (Maule et al., 2000) or has diminished cognitive ability (Fallon et al., 2014; Figure 2, right, bottom). As one would expect, suicidal people generate fewer solutions when presented with problems (Evans et al., 1992; Pollock & Williams, 2004). When a smaller set is sampled from a universe of possible solutions under these constraints, the expected maximum value of the best solution in the set is also lower, increasing the probability that suicide will be selected in the final choice process. This narrowing of the consideration set is exacerbated by Pavlovian biases, resulting in a consideration set, which is both undersized and congruent with depressed mood.
Crises are often not as pressing as people believe them to be. Suicidal people experience a heightened sense of urgency (Dombrovski et al., 2019; Klonsky & May, 2010) and endorse favoring the first solution that comes to mind. Within the decision process framework, urgency means assigning inaction a low relative value. Negative urgency (e.g., “I would do anything to stop feeling this way”) is related to the personality dimensions of neuroticism and impulsivity discussed in Section 2 and appears to invigorate behaviors that are disadvantageous in the long term such as drug use and disordered eating (Racine et al., 2013; Torres et al., 2013). Curiously, self-reported urgency does not distinguish suicide attempters from ideators (Dombrovski et al., 2019; Klonsky & May, 2010), suggesting that it operates by accelerating decision-making and distorting consideration sets, but does not necessarily favor the final selection of suicide over other options.

A large part of the process of planning the suicidal act is predecisional. People who end up attempting suicide have often chosen the method and even location days or weeks prior to deciding to go through with the plan (Millner et al., 2017). In our research studies, we see many well-planned suicide attempts (Dombrovski et al., 2011, 2019; Dombrovski, Szanto, et al., 2008; Szanto et al., 2018, 2020). In agreement with Millner, however, we also see people seeking help after they initiated planning (e.g., final arrangements, hoarding medications), indicating that the decision had not been made. Similar predecisional constraints are imposed before one decides to purchase durable goods, such as a new car: even before the need arises, people may know that their next car will be German or an SUV. Unsurprisingly, people with more predecisional constraints end up reporting higher confidence in their choice (Punj & Brookes, 2002). It follows that predecisional suicide planning would increase the odds of eventual suicide by rendering this option more certain and reducing the need for cognitively taxing simulations discussed in the next section. Indeed, as we discuss below, the perceived simplicity and feasibility of suicide distinguishes it from alternative solutions, which are likely to entail multiple steps with uncertain outcomes.

In summary, the development of suicidal ideation can be understood in terms of selection of options into a consideration set based on precomputed rather than situation-specific values. These values, furthermore, are biased by one’s...
affective states. Consideration sets composed under cognitive constraints are obviously less likely to include superior but harder-to-find situation-appropriate alternative solutions (e.g., hiring a professional mediator or divorce lawyer who handles all contact with the spouse, serving as an emotional buffer and advocate), increasing the likelihood that suicide will be eventually selected. These basic cognitive mechanisms are well understood, but their role in suicide is hypothetical, and experimental studies of consideration set construction in people vulnerable to suicide are needed.

4.5 | Learning and simulation

4.5.1 | Reinforcement learning (Box 2)

RL models include a value function and a policy. The learning process uses experience to estimate values of states of the environment, stimuli, or actions and is subject to cognitive constraints. The policy selects among available actions, in this case suicidal behavior versus available alternatives, based on their values. Human learning likely relies on several parallel algorithms. In canonical RL, one uses the delta rule (Box 2) to learn and cache values of each possible state of the environment (Bush & Mosteller, 1951; Rescorla & Wagner, 1972; Watkins & Dayan, 1992). This simple form of learning, thought to depend on the basal ganglia-mesostriatal circuit, is computationally light but too inflexible in complex, changing environments. It is often referred to as model-free learning, in contrast to computationally taxing but often more efficient model-based learning, where the agent uses a model of the world to simulate possible paths and downstream outcomes of a decision (Daw et al., 2011), a process thought to depend on the prefrontal cortex (Otto, Gershman, et al., 2013; Smittenaar et al., 2013). Some of the more sophisticated aspects of human decision-making previously unexplained by model-free RL—generalizing across contexts, judging what information is currently relevant, breaking down big goals into subgoals—depend on interactions of canonical RL with cognitive control processes subserved by fronto-parietal, cingulo-operandal networks (Botvinick, 2012; J. F. Cavanagh & Frank, 2014; Collins et al., 2017; Otto, Gershman, et al., 2013; Rmus et al., 2020). On the other hand, cognitive maps that support learning in physical and abstract spaces depend on the hippocampus (Dombrovski et al., 2020; Mattar & Daw, 2018; Miller et al., 2017; Vikbladh et al., 2019). Once the values are estimated, the agent makes a choice, a process discussed in the next section (Section 4.6).

4.5.2 | Learning deficits that may propel the suicidal crisis

Clinical theorists see the narrow, short-sighted cognitive perspective as a key feature of the suicidal crisis (Baumeister, 1990; Shneidman et al., 1970; Weiss, 1957), and in our clinical experience people in a suicidal crisis often seem unaware of the tragic consequences of suicide for their families and of potentially effective alternative solutions. What might account for this misprediction? Probably not a deficit in model-free learning since, as Skinner noted, suicide as an operant cannot be strengthened.1 In our framework, the suicidal crisis is propelled by a failure to integrate new experiences and simulate possible outcomes, explaining the outsize effect of seemingly modest stressors, catastrophizing, and the inability to find solutions. The deficit in question must undermine people's sensitivity to the downstream consequences of their actions, which can involve two distinct mechanisms, namely transition revaluation and reward revaluation (Momennejad et al., 2017).

Transition revaluation is behavioral sensitivity to causal links between upstream actions and their downstream consequences, dependent on the posterior hippocampus/entorhinal cortex, dorsomedial striatum, and the orbitofrontal cortex (as shown by classical contingency degradation experiments; Corbit & Balleine, 2000; Jackson et al., 2016; Lex & Hauber, 2010). It is likely that in complex, dynamic environments the ability to simulate multiple trajectories depends on the dorsolateral prefrontal and cingulate cortices (Collins et al., 2017; Wang et al., 2018). Such a deficit in representing causal chains of events can propel the suicidal crisis by undermining mental simulations that evaluate alternative solutions. The relative value of suicide then increases because of its perceived simplicity and certainty, often enhanced by pre-decisional planning (discussed in Section 4.2, Millner et al., 2017). Reward revaluation involves updating the value of upstream actions based on changes in the value of downstream outcomes, shown in outcome devaluation experiments to depend on the orbitofrontal cortex and basolateral amygdala (Balleine et al., 2003; Corbit & Balleine, 2005; Panayi & Killcross, 2018). A deficit in reward revaluation can undermine deterrents such as foregoing future joys and hurting one's family because when these outcomes are considered, their emotional value does not serve
to update the perceived value of suicide. Simply speaking, people may lack a visceral sense of how the consequences of suicide or its alternatives would feel. Indeed, suicide attempters display blunted affective forecasting for future positive events (Marroquín et al., 2013).

### 4.5.3 Learning in suicidal behavior

We now turn to the question whether suicidal people fail to learn and simulate the downstream consequences of their actions. Such a failure may result from impaired transition evaluation or reward revaluation. Empirically, suicide attempters show deficits on probabilistic learning tasks (Dombrovski et al., 2010, 2019) and in some (McGirr et al., 2012), but not all, studies of deterministic learning using the Wisconsin Card Sort (WCST) (meta-analysis: Richard-Devantoy, Berlim, et al., 2014). In our recent two-sample study \((n = 135/125, n_{\text{attempters}} = 54/39; \text{Table 3, Learning})\) using a three-armed bandit task, suicide attempters with mid-life and late-life depression (vs. suicide ideators, nonsuicidal patients with depression, and healthy control participants) failed to encode one-back reinforcement, as seen in analyses of choices as well as decision times. Learning deficits scaled with attempt lethality (Dombrovski et al., 2019). These were partially explained by poor cognitive control among suicide attempters, particularly those with high-lethality attempts, echoing the pattern of group differences in our preceding studies of the WCST and Stroop inhibition (McGirr et al., 2012; Richard-Devantoy, Szanto, et al., 2014) and the broader literature on Stroop deficits in attempted suicide (Richard-Devantoy, Berlim, et al., 2014). Stroop and WCST performance depends on the broad cognitive control network, encompassing the dorsal ACC (Gläscher et al., 2012) and the lateral PFC (Yuan & Raz, 2014), which as we noted earlier supports the ability to represent and simulate action sequences (Collins et al., 2017; Wang et al., 2018). To conclude, although the studies reviewed here did not directly test transition revaluation, co-occurring and associated deficits in learning and cognitive control are generally consistent with impaired transition revaluation, undermining one’s ability to simulate action sequences necessary for alternative solutions. To the best of our knowledge, no studies have tested the reward revaluation hypothesis, that is the inability to update the value of upstream actions based on dynamic values of downstream outcomes undermines deterrents to suicide.

### TABLE 3 Replicable deficits in learning and value-based choice processes in attempted suicide (Dombrovski et al., 2019)

| Decision deficits on a three-armed bandit task | Reference group | Groups differing from the reference group |
|-----------------------------------------------|-----------------|----------------------------------------|
| Learning                                      | Suicide attempters | Controls (C), depressed (D), ideators (I) | C, I, D |
|                                               | High-lethality attempters | C, D, I | C, I |
| Exaggerated post-reward decision time slowinga | Suicide attempters | C, D, I | C, I, C, D, I, low-lethality attempters (LL) |
|                                               | High-lethality attempters | C, D, I, LL | C, I, LL |
| Choice                                        | Suicide attempters | C, I | C, D, I |
| Poor discrimination among close-valued options | Suicide attempters | C, D, I | C, D, I |

Note: C, healthy controls; D, nonsuicidal depressed; I, suicide ideators; LL, low-lethality suicide attempters. Behavior was analyzed with mixed-effects binary logistic hierarchical model. Decision times analyses employed a reinforcement learning model inverted using empirical Bayes combined with a linear mixed-effects model. There were no differences between high- and low-lethality attempters in choice analyses.

aGroup differences in learning were partially explained by cognitive control (modified from Dombrovski et al., 2019).
4.6 Choice processes

We know relatively little about the transition from contemplation to suicidal action (Franklin et al., 2017; Klonsky et al., 2018; Saffer & Klonsky, 2018). A person choosing between suicide and its alternatives engages in an implicit comparison of their expected values. The rarity of suicide suggests that its value is dominated by those of alternatives for the overwhelming majority of people, most of the time. If so, greater randomness in the value-based choice process would increase the rate at which suicide is selected. The notion that value-based choice is disrupted in people at risk for suicide is supported by the IGT findings reviewed earlier and, indirectly, by post-mortem and early autoradiographic findings of alterations in the ventromedial prefrontal cortex (Amen et al., 2009; Arango et al., 1997; Oquendo et al., 2003) subserving construction and comparison of values in both learning and revealed preference paradigms (Ballesta et al., 2020; Bartra et al., 2013; Chase et al., 2015). The idea that suicide reflects a stochastic choice is not new. Beskow et al. (1994), for example, conceptualized suicide as a "psychological accident." Drawing on demand-capability models from the occupational safety literature, Beskow argued that when demands (stressors or access to lethal means) exceed capability (coping), vulnerable individuals respond with stochastic ("chaotic") suicidal acts.

4.6.1 Formal models of the choice process

RL traditionally simplifies the choice process by entering action values into a function, most commonly the softmax (Box 2), which returns the probabilities with which actions are chosen. In real life, however, the process of deciding between suicide, inaction, and alternative solutions unfolds in real time over hours, days, or weeks (Bryan et al., 2019; Millner et al., 2017). Clearly, not only what is eventually chosen matters, but also when. The temporal dynamics of the choice process can be described by sequential sampling models, such as drift-diffusion (DDM; Pedersen et al., 2017; Ratcliff & Rouder, 1998). In this dynamic account, the agent drifts stochastically between alternative actions (or an action and inaction) depending on their perceived values shaped by ongoing experience. Whereas the choices of DDM mimic the softmax (Baldassi et al., 2020), its parameters additionally control decision times. Here, we discuss a binary choice case, but DDM can be extended to multi-alternative choice (Krajbich & Rangel, 2011).

One substantively interesting parameter in DDM is boundary separation or response caution trading off accuracy versus speed (Hedge et al., 2020; Shapiro & Huang-Pollock, 2019). Tighter boundaries produce faster and less accurate decisions and in theory may result in suicide being preferred to superior alternatives. Boundary separation, however, appears to be task-specific and may not constitute a general cognitive trait (Schubert et al., 2016). Another parameter, drift or evidence accumulation rate, is more likely to represent a true trait, although psychometric data come mostly from “cold” cognitive tasks as opposed to value-based decision-making, and link drift rate to working memory and fluid intelligence (Schmeidek et al., 2007; Schubert et al., 2016). During value-based decision-making under uncertainty, drift rate scales with option values (Pedersen et al., 2017; Shahar et al., 2019), with decisions becoming slower and less value-sensitive at lower drift rates. In the case of the suicidal crisis, a person diffusing between suicide and alternatives (or inaction), tighter boundaries would result in impulsive suicidal acts whereas lower drift would manifest in vacillation sometimes ending in a suicidal act. However, the task-specific nature of boundary parameters means that a decision boundary operating during the suicidal process may be difficult to recover in the lab. In summary, low drift rates best formalize the stochastic choice hypothesis, leading to the prediction that affected individuals will make slow and inconsistent choices, a tendency that can prove fatal when suicide is in the consideration set. Drift rate, of course, may be low and choices may be stochastic only because values were learned poorly. Thus, if the abnormality lies specifically and exclusively in the choice process, stochastic choice in general and low drift rate should be seen not only on learning tasks (such as armed bandits), but also on value-based decision tasks that do not involve learning.

One objection to the stochastic choice hypothesis is that high-stakes choices are more deliberate and thus more predictable. Indeed, this hypothesis cannot explain the most premeditated suicides. Yet, interestingly, humans tend not to calibrate their deliberation time according to the value at stake, rushing important decisions (Oud et al., 2016). Thus, the stochastic component diminishes, but not as much as it should, even for life-and-death decisions, as seen from medical errors, judicial errors in capital cases, or misguided (in hindsight) declarations of war.
Evidence of stochastic value-based choice in suicidal behavior

As we noted earlier, the IGT studies are consistent with the stochastic choice hypothesis (meta-analyses: Richard-Devantoy, Berlim, et al., 2014), but inconclusive because the choice process was not separated from learning, by design or analytically. This ambiguity can be resolved by data from gambling tasks that do not involve learning, such as the Cambridge Gamble and Game of Dice Tasks (CGT and GDT). Our study using the Cambridge Gamble found that older depressed suicide attempters made more stochastic choices than suicide ideators, nonsuicidal depressed patients and healthy controls (Clark et al., 2011), and in a later study stochastic choices on the CGT predicted blunted vmPFC value signals in the same patient cohort (Dombrovski et al., 2013). A subsequent larger study of young adults confirmed these behavioral findings (Chamberlain et al., 2013). While a later small CGT study in adolescent suicide attempters was inconclusive (Ackerman et al., 2014), a meta-analysis (Perrain et al., 2021) of the three studies confirmed the findings of CGT impairment in suicide attempters versus patient controls. Furthermore, in a study using the GDT adolescent suicide attempters made choices that were more risk-seeking and/or stochastic (Deisenhammer et al., 2018).

Our recent study of two depression samples using a three-armed bandit paradigm sought to separate choice processes from learning using RL modeling. We observed that suicide attempters (vs. all comparison groups) exhibited longer decision times for choices where options differed little in their expected value and made win-switch choices suggestive of failure to exploit the best available value (Dombrovski et al., 2019), Table 3, Choice. Excessive—rather than lacking—decision time modulation by long-term values could not be explained by random fluctuations, poor effort, or distraction. These effects were robust to controlling for cognitive control, estimated premorbid IQ and general cognitive ability, as well as medication exposure and possible brain damage from suicide attempts. Since participants chose between three options, it was not straightforward to fit a DDM model, however longer decision times for choices between close-valued options are broadly consistent with low drift rate. Overall, findings from IGT, non-learning gambling tasks and multi-armed bandit studies are consistent with the hypothesis that a tendency toward stochastic value-based choices facilitates the progression from contemplating to attempting suicide. The tendency toward stochastic choices in suicide attempters appears to be independent of deficits in learning or cognitive control. Existing evidence does not permit direct inference about parameters that control dynamic choice processes in real time but is broadly consistent with a low DDM drift rate.

Intertemporal preferences: myopic or stochastic?

One may find Cicero’s perspective on suicide naïve, since the evaluation of one’s future depends on psychological state and their ability to resolve problems. Yet adopting this normative position for a moment, we can view suicide as a choice that reflects a negative expected value of one’s life. As this value is integrated over a sequence of current and future (“... in prospect”) experiences, preferences such as the weight one places on immediate versus delayed outcomes may determine whether the expected value of one’s entire life is positive or negative. Indeed, short-term crises often precipitate suicide attempts, suggesting that people vulnerable to suicidal behavior may be myopically present-focused. This myopia may be exaggerated by negative affective states, which may truncate future evaluation through a Pavlovian mechanism discussed above (4.5.4). Early studies have associated an exaggerated preference for immediate over delayed rewards—delay discounting, broadly associated with self-reported impulsivity (Kirby & Finch, 2010)—with some forms of suicidal behavior, including low-lethality, poorly planned suicide attempts in late-life depression (Dombrovski et al., 2011, 2012), suicide attempts among people with addiction (Liu et al., 2012) and among adolescents (Dougherty et al., 2009; Mathias et al., 2011). Other studies of adolescent (Bridge et al., 2015) or adult suicide attempters (Millner et al., 2018) yielded no evidence of increased delay discounting. One explanation for this lack of agreement between studies is that estimated discount rates in suicidal people and in clinical populations in general may be unreliable because of choice inconsistency or stochasticity (Bartolomeo et al., 2020; Coelho et al., 2017; Lindbergh et al., 2014a, 2014b). Thus, the inability to integrate disparate attributes of options (in this case, reward magnitude and delay) manifesting in inconsistency may underlie suicidal people’s choice of immediate relief at the expense of the future and in spite of deterrents.

Cognitive diathesis to suicide across the lifespan

It is likely that the nature of the suicidal diathesis changes across the lifespan. Early-onset (vs. late-onset, typically after age 40 or 50) suicidal behavior is more likely to be associated with a greater burden of psychopathology, personality
pathology in general (Mittendorfer-Rutz et al., 2008; Szücs et al., 2020), and externalizing or impulsive-aggressive traits in particular (McGirr et al., 2008). The same is true for early environment factors, including abuse and neglect and social transmission of suicidal behavior. By contrast, natural cognitive variability produced by neurodegenerative and vascular processes in old age and the prominence of cognitive deficits in late- versus early-onset depression (meta-analysis: Bora et al., 2013) suggest that the cognitive diathesis may play a comparatively greater role in late-onset cases. Our prospective study of attempted suicide in late life supports this notion, finding that poor cognitive control predicts high-lethality suicide attempts whereas personality factors predict low-lethality attempts (Szanto et al., 2018, 2020). Our prior retrospective case–control study found a similar relationship between cognitive impairment and attempted suicide (Dombrovski, Butters, et al., 2008; Gujral et al., 2014, 2016). These data suggest that many late-life suicides occur in a prodrome of dementia. This is a testable hypothesis, but one complication is that people with advanced dementia cannot plan and, at some point, even execute suicidal acts. In our clinical experience, they tend to forget their earlier suicidal plans; we refer to this protective influence as the Lethe effect.3 The competing risk of natural death further masks suicide risk. Due to these effects any window of risk is limited to prodromal and early disease. Mixed findings concerning the relationship between dementia and suicide (reviewed in Draper, 2015) are thus not surprising, although patients with behavioral problems triggering a hospitalization appear to be at higher risk (Erlangsen et al., 2008). Along similar lines, a recent large Danish population study found increased incidence of suicide in Huntington’s disease, while Alzheimer’s disease was paradoxically protective (Erlangsen et al., 2020). Interestingly, a national Taiwanese study may have circumvented the Lethe effect by examining the risk of future dementia in older suicide attempters versus non-attempters: dementia incidence was higher in attempters even after controlling for psychopathology confounds (Tu et al., 2016).

One disease of theoretical interest to us is the behavioral variant of frontotemporal lobar degeneration (bvFTD) as in early stages it selectively affects the ventromedial prefrontal cortex (vmPFC), a key region implicated in value-based choice (Section 4.6). We have previously hypothesized that vmPFC is a primary site of neural abnormalities underlying decision deficits that predispose to suicidal behavior (Clark et al., 2011; Dombrovski et al., 2013). A study of 56,000 US veterans found that the prevalence of suicidal behavior (defined as plan or attempt) was 0.42% in bvFTD, compared to 0.06% in Alzheimer’s disease, 0.20% in vascular dementia, and 0.03% in Lewy body dementia, with a similar pattern for suicidal ideation (Lai et al., 2018). As we discussed earlier (Section 4.6.2), the vmPFC/OFC is necessary for two forms of action: bideo decisions (Section 4.6). We have previously hypothesized that vmPFC is a primary site of neural abnormalities underlying decision deficits that predispose to suicidal behavior (Clark et al., 2011; Dombrovski et al., 2013). A study of 56,000 US veterans found that the prevalence of suicidal behavior (defined as plan or attempt) was 0.42% in bvFTD, compared to 0.06% in Alzheimer’s disease, 0.20% in vascular dementia, and 0.03% in Lewy body dementia, with a similar pattern for suicidal ideation (Lai et al., 2018). As we discussed earlier (Section 4.6.2), the vmPFC/OFC is necessary for two forms of model-based learning, transition revaluation and reward revaluation, and thus people with bvFTD may be unable to simulate the action sequences involved in alternative solutions or weight the impact of their suicide on others. vmPFC lesions also result in more random value-based choice and may thus cause suicide to be selected even in the presence of alternatives, which are known to be superior.

4.8 | Going further: cognitive overload and resource-optimal decision-making

One aspect of the suicidal crisis that paradigms discussed so far do not fully capture is cognitive overload. Patients report feeling lost and not knowing where to start, and clinical theorists speak of concreteness, lack of distal goals, and passivity (Baumeister, 1990). As Beskow et al. (1994) and Pollock and Williams (2004) noted, the suicidal crisis escalates when cognitive demands exceed one’s capacity. In the uncertainty of a crisis, the need to appraise many options at once imposes a high information load and degrades the quality of decision-making. Successful decision-making under high information load relies not merely on working memory capacity, but on resource-rational strategies, which balance accuracy with time and cognitive effort (Griffiths et al., 2015). In the absence of such strategies, the quality of decision-making deteriorates catastrophically once capacity is exceeded, manifesting in stochastic choices. More specifically, the value of computation theory suggests that we allocate cognitive resources to achieve the best tradeoff between the expected utility of the solution and the opportunity cost of time required for computation (Lieder et al., 2014; Otto & Daw, 2019; Shenhav et al., 2017). Compression is one such resource-rational strategy dealing with the problem of multiple options of uncertain value, which imposes a high information load. We found that people who successfully transition from exploring multiple options to exploiting the most advantageous ones selectively remember the values of the best options and forget the rest, compressing the information that needs to be remembered (Hallquist & Dombrovski, 2019). We have preliminary evidence that older depressed suicide attempters make excessively stochastic choices when faced with multiple options and fail to engage in information-compressing selective maintenance. Future studies will need to examine how people vulnerable to suicidal behavior deal with information overload and constraints such as time pressure (Otto & Daw, 2019) during decision-making.
**Conclusion and future directions**

Suicide may be viewed as an unfortunate outcome of failures in one or more decision processes, occurring when a person’s capacity to search for options, learn, simulate possible futures, and make good choices is exceeded by the demands of a crisis. We divide the suicidal process into three stages: (i) the construction of a consideration set of solutions based on their precomputed values in response to urgent demands, (ii) learning and simulation evaluating the alternative solutions, and (iii) the choice process based on updated values. Broadly speaking, decision processes in many suicidal people are subject to crippling cognitive constraints of limited cognitive capacity and particularly impaired cognitive control. Suicidal ideas can be seen as an element in impoverished consideration sets constructed under cognitive constraints. This view is supported by considerable indirect evidence, but experimental studies are lacking. On the other hand, a sizeable body of evidence supports impaired learning in suicide attempters. We think that a deficit in transition revaluation— or the ability to simulate possible futures based on one’s model of the world—is most likely to undermine alternative solutions in a suicidal crisis, and experiments testing this idea directly are needed. The hypothesis that rests on perhaps the strongest evidence is that the selection of suicide over alternatives is facilitated by a choice process undermined by randomness. Studies using gambling tasks and armed bandits support this claim. Little is known, however, about real-time dynamics of choice processes in suicidal people, and experimental studies dissecting these dynamics are needed. Table 4 gives some examples of critical studies needed to test these hypotheses. To conclude, the decision process investigative framework allows us to examine the neurocognitive mechanisms that underlie various stages of the suicidal crisis, bridging clinical theory and neurobiology of suicidal behavior.

**ACKNOWLEDGMENT**

We thank Aliona Tsypes for her comments on the manuscript.

**CONFLICT OF INTEREST**

The authors have declared no conflicts of interest for this article.

**DATA AVAILABILITY STATEMENT**

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

**AUTHOR CONTRIBUTIONS**

Alexandre Y. Dombrovski: Conceptualization; funding acquisition; writing-original draft; writing-review & editing. Michael N. Hallquist: Conceptualization; funding acquisition; writing-original draft; writing-review & editing.
ENDNOTES

1 One important exception are recurrent suicidal acts which lead to hospitalization.

2 The stochastic choice account probably cannot explain the most determined suicidal acts.

3 Of the two rivers bounding the Greek underworld, dementia forces you to drink from the Lethe before you cross the Styx.

RELATED WIREs ARTICLES

Model-based approaches to neuroimaging: combining reinforcement learning theory with fMRI data
Computational perspectives on cognitive development

REFERENCES

Ackerman, J. P., McBee-Strayer, S. M., Mendoza, K., Stevens, J., Sheffall, A. H., Campo, J. V., & Bridge, J. A. (2014). Risk-sensitive decision-making deficit in adolescent suicide attempters. Journal of Child and Adolescent Psychopharmacology, 25(2), 109–113. https://doi.org/10.1089/cap.2014.0041

Aghbayewa, M. O., Marion, S. A., & Wiggins, S. (1998). Socioeconomic factors associated with suicide in elderly populations in British Columbia: An 11-year review. Canadian Journal of Psychiatry, 43(8), 829–836.

Agerbo, E., Nordentoft, M., & Mortensen, P. B. (2002). Familial, psychiatric, and socioeconomic risk factors for suicide in young people: Nested case-control study. BMJ: British Medical Journal, 325(7355), 74.

Allen, K. J. D., Bozayk, M. L., & Edenbaum, E. R. (2019). Neurocognition and suicide risk in adults. Current Behavioral Neuroscience Reports, 6(4), 151–165. https://doi.org/10.1007/s40473-019-00189-y

Allen, T. A., Dombrovski, A. Y., Soloff, P. H., & Hallquist, M. N. (2020). Borderline personality disorder: Stress reactivity or stress generation? A prospective dimensional study. Psychological Medicine, 1–8. https://doi.org/10.1017/S003329172000255X

Allen, T. A., Schreiber, A. M., Hall, N. T., & Hallquist, M. N. (2020). From description to explanation: Integrating across multiple levels of analysis to inform neuroscientific accounts of dimensional personality pathology. Journal of Personality Disorders, 34(5), 650–676. https://doi.org/10.1521/pedi.2020.34.5.650

Allman, M. J., DeLeon, I. G., Cataldo, M. F., Holland, P. C., & Johnson, A. W. (2010). Learning processes affecting human decision making: An assessment of reinforcer-selective Pavlovian-to-instrumental transfer following reinforcer devaluation. Journal of Experimental Psychology: Animal Behavior Processes, 36(3), 402–408. https://doi.org/10.1037/a0017876

Amen, D. G., Prunella, J. R., Fallon, J. H., Amen, B., & Hanks, C. (2009). A comparative analysis of completed suicide using high resolution brain SPECT imaging. The Journal of Neuropsychiatry and Clinical Neurosciences, 21(4), 430–439. https://doi.org/10.1176/jnp.2009.21.4.430

Andersson, L., Allebeck, P., Gustafsson, J. E., & Gunnell, D. (2008). Association of IQ scores and school achievement with suicide in a 40-year follow-up of a Swedish cohort. Acta Psychiatrica Scandinavica, 118(2), 99–105. https://doi.org/10.1111/j.1600-0447.2008.01171.x

Anestis, M. D., Soberay, K. A., Gutierrez, P. M., Hernández, T. D., & Joiner, T. E. (2014). Reconsidering the link between impulsivity and suicidal behavior. Personality and Social Psychology Review, 18(4), 366–386. https://doi.org/10.1177/1088868314535988

Arango, V., Underwood, M. D., & Mann, J. J. (1997). Postmortem findings in suicide victims: Implications for in vivo imaging studies. Annals of the New York Academy of Sciences, 836, 269–287.

Baldassi, C., Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., & Pirazzini, M. (2020). A behavioral characterization of the drift diffusion model and its multialternative extension for choice under time pressure. Management Science, 66, 5075–5093. https://doi.org/10.1287/mnsc.2019.3475

Ballard, E. D., Voort, J. L. V., Luckenbaugh, D. A., Machado-Vieira, R., Tohen, M., & Zarate, C. A. (2016). Acute risk factors for suicide attempts and death: Prospective findings from the STEP-BD study. Bipolar Disorders, 18(4), 363–372. https://doi.org/10.1111/bdi.12397

Balleine, B. W., Killcross, A. S., & Dickinson, A. (2003). The effect of lesions of the basolateral amygdala on instrumental conditioning. Journal of Neuroscience, 23(2), 666–675. https://doi.org/10.1523/JNEUROSCI.23-02-00666.2003

Ballesta, S., Shi, W., Conen, K. E., & Padoa-Schioppa, C. (2020). Values encoded in orbitofrontal cortex are causally related to economic choices. Nature, 588, 450–453. https://doi.org/10.1038/s41586-020-2880-x

Bartolomeo, L. A., Chapman, H. C., Raugh, I. M., & Strauss, G. P. (2020). Delaying discounting in youth at clinical high-risk for psychosis and adults with schizophrenia. Psychological Medicine, 1–8. https://doi.org/10.1017/S0033291720000677

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. NeuroImage, 76, 412–427. https://doi.org/10.1016/j.neuroimage.2013.02.063

Batty, G. D., Whitley, E., Deary, I. J., Gale, C. R., Tynelius, P., & Rasmussen, F. (2010). Psychosis alters association between IQ and future risk of attempted suicide: Cohort study of 1 109 475 Swedish men. BMJ, 340, c2506. https://doi.org/10.1136/bmj.c2506

Baumeister, R. F. (1990). Suicide as escape from self. Psychological Review, 97(1), 90–113.
Shneidman, E. S., Farberow, N. L., & Litman, R. E. (1970). The psychology of suicide. https://repository.library.georgetown.edu/handle/10822/763228

Silverman, M. H., Wilson, S., Ramsay, I. S., Hunt, R. H., Thomas, K. M., Krueger, R. F., & Iacono, W. G. (2019). Trait neuroticism and emotion neurocircuitry: Functional magnetic resonance imaging evidence for a failure in emotion regulation. Development and Psychopathology, 31(3), 1085–1099. https://doi.org/10.1017/S0954579419000610

Smittenaar, P., FitzGerald, T. H. B., Romei, V., Wright, N. D., & Dolan, R. J. (2013). Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. Neuron, 80(4), 914–919. https://doi.org/10.1016/j.neuron.2013.08.009

Snyder, H. R. (2013). Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: A meta-analysis and review. Psychological Bulletin, 139(1), 81–132. https://doi.org/10.1037/a0028727

Sokolowski, M., & Wasserman, D. (2020). Genetic origins of suicidality? A synopsis of genes in suicidal behaviours, with regard to evidence diversity, disorder specificity and neurodevelopmental brain transcriptomics. European Neuropsychopharmacology, 37, 1–11. https://doi.org/10.1016/j.euroneuro.2020.06.002

Steele, I. H., Thrower, N., Noroian, P., & Saleh, F. M. (2018). Understanding suicide across the lifespan: A United States perspective of suicide risk factors, assessment & management. Journal of Forensic Sciences, 63(1), 162–171. https://doi.org/10.1111/1556-4029.13519

Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction. MIT Press.

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning, second edition: An introduction. MIT Press.

Szanto, K., Galfalvy, H., Kenneally, L., Almasi, R., & Dombrovski, A. Y. (2020). Predictors of serious suicidal behavior in late-life depression. European Neuropsychopharmacology, 30, 85–98. https://doi.org/10.1016/j.euroneuro.2020.06.005

Szanto, K., Galfalvy, H., Vanyukov, P. M., Keilp, J. G., & Dombrovski, A. Y. (2018). Pathways to late-life suicidal behavior: Cluster analysis and predictive validation of suicidal behavior in a sample of older adults with major depression. The Journal of Clinical Psychiatry, 79(2), 17m11611. https://doi.org/10.4088/JCP.17m11611

Szücs, A., Szanto, K., Wright, A. G. C., & Dombrovski, A. Y. (2020). Personality of late- and early-onset elderly suicide attempters. International Journal of Geriatric Psychiatry, 35(4), 384–395. https://doi.org/10.1002/gps.5254

Torres, A., Catena, A., Megias, A., Maldonado, A., Cándido, A., Verdejo-García, A., & Perales, J. C. (2013). Emotional and non-emotional pathways to impulsive behavior and addiction. Frontiers in Human Neuroscience, 7, 43. https://doi.org/10.3389/fnhum.2013.00043

Tu, Y.-A., Chen, M.-H., Tsai, C.-F., Su, T.-P., Bai, Y.-M., Li, C.-T., Lin, W.-C., Liu, M.-E., Pan, T.-L., Chen, T.-J., & Tsai, S.-J. (2016). Geriatric suicide attempt and risk of subsequent dementia: A nationwide longitudinal follow-up study in Taiwan. The American Journal of Geriatric Psychiatry, 24, 1211–1218. https://doi.org/10.1016/j.jagp.2016.08.016

Turecki, G., Brent, D. A., Gunnell, D., Oquendo, M. A., Pirkis, J., & Stanley, B. H. (2019). Suicide and suicide risk. Nature Reviews. Disease Primers, 5(1), 74. https://doi.org/10.1038/s41572-018-0121-0

Vaghl, M. M., Cardinal, R. N., Aerzigis-Schoute, A. M., Fineberg, N. A., Sule, A., & Robbins, T. W. (2019). Action-outcome knowledge dissociates from behavior in obsessive-compulsive disorder following contingency degradation. Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 4(2), 200–209. https://doi.org/10.1016/j.bpsc.2018.09.014

Van Orden, K. A., Witte, T. K., Cukrowicz, K. C., Braithwaite, S. R., Selby, E. A., & Joiner, T. E., Jr. (2010). The interpersonal theory of suicide. Psychological Review, 117(2), 575–600. https://doi.org/10.1037/a0018697

Verona, E., Sachs-Ericsson, N., & Joiner, T. E. (2004). Suicide attempts associated with externalizing psychopathology in an epidemiological sample. American Journal of Psychiatry, 161(3), 444–451. https://doi.org/10.1176/appi.ajp.161.3.444

Vijayakumar, L., Kumar, M. S., & Vijayakumar, V. (2011). Substance use and suicide. Current Opinion in Psychiatry, 24(3), 197–202. https://doi.org/10.1097/YCO.0b013e3283459242

Vikbladh, O. M., Meager, M. R., King, J., Blackmon, K., Devinsky, O., Shohamy, D., Burgess, N., & Daw, N. D. (2019). Hippocampal contributions to model-based planning and spatial memory. Neuron, 102, 683–693.e4. https://doi.org/10.1016/j.neuron.2019.02.014

Voracek, M. (2004). National intelligence and suicide rate: An ecological study of 85 countries. Personality and Individual Differences, 37(3), 543–553.

Voracek, M. (2006a). Regional intelligence and suicide rate in Denmark. Psychological Reports, 98(3), 671–674.

Voracek, M. (2006b). Regional intelligence and suicide rate in Germany. Perceptual and Motor Skills, 103(2), 639–642.

Waern, M., Rubenowitz, E., Runeson, B., Skoog, I., Wilkinson, K., & Allebeck, P. (2002). Burden of illness and suicide in elderly people: Case–control study. BMJ, 324(7350), 1355.

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., Hassabis, D., & Botvinick, M. (2018). Prefrontal cortex as a meta-reinforcement learning system. Nature Neuroscience, 21(6), 860–868. https://doi.org/10.1038/s41593-018-0147-8

Watkins, C. J. C. H., & Dayan, P. (1992). Q-learning. Machine Learning, 8(3–4), 279–292. https://doi.org/10.1007/BF00992698

Webb, R. T., Längström, N., Runeson, B., Lichtenstein, P., & Fazel, S. (2011). Violent offending and IQ level as predictors of suicide in schizophrenia: National cohort study. Schizophrenia Research, 130(1), 143–147. https://doi.org/10.1016/j.schres.2011.04.031

Weiss, J. M. A. (1957). The gamble with death in attempted suicide. Psychiatry, 20(1), 17–25. https://doi.org/10.1080/00332747.1957.11023072

Wenzel, A., Berchick, E. R., Tenhave, T., Halberstadt, S., Brown, G. K., & Beck, A. T. (2011). Predictors of suicide relative to other deaths in patients with suicide attempts and suicide ideation: A 30-year prospective study. Journal of Affective Disorders, 132(3), 375–382. https://doi.org/10.1016/j.jad.2011.03.006

WHO Mental Health. (2016). WHO | Suicide data. World Health Organization http://www.who.int/mental_health/prevention/suicide/suicidereprev/en/

Williams, J. M. G., Crane, C., Barnhofer, T., & Duggan, D. (2005). Psychology and suicidal behaviour: Elaborating the entrapment model. In K. Hawton (Ed.), Prevention and treatment of suicidal behaviour: From science to practice (1st ed., pp. 71–89). Oxford University Press.
Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General, 143*(6), 2074–2081.

Wong, P. W., Cheung, D. Y., Conner, K. R., Conwell, Y., & Yip, P. S. (2010). Gambling and completed suicide in Hong Kong: A review of coroner court files. *The Primary Care Companion to The Journal of Clinical Psychiatry, 12*(6), e1–e7. https://doi.org/10.4088/PCC.09m00932blu

World Health Organization. (1998). *Primary prevention of mental, neurological and psychosocial disorders*. Author.

Wu, C. M., Schulz, E., Speekenbrink, M., Nelson, J. D., & Meder, B. (2018). Generalization guides human exploration in vast decision spaces. *Nature Human Behaviour, 2*(12), 915–924. https://doi.org/10.1038/s41562-018-0467-4

Yip, P. S. F., Caine, E., Yousuf, S., Chang, S.-S., Wu, K. C.-C., & Chen, Y.-Y. (2012). Means restriction for suicide prevention. *Lancet (London, England), 379*(9834), 2393–2399. https://doi.org/10.1016/S0140-6736(12)60521-2

Yuan, P., & Raz, N. (2014). Prefrontal cortex and executive functions in healthy adults: A meta-analysis of structural neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 42*, 180–192. https://doi.org/10.1016/j.neubiorev.2014.02.005

Zatti, C., Rosa, V., Barros, A., Valdivia, L., Calegaro, V. C., Freitas, L. H., Ceresér, K. M. M., da Rocha, N. S., Bastos, A. G., & Schuch, F. B. (2017). Childhood trauma and suicide attempt: A meta-analysis of longitudinal studies from the last decade. *Psychiatry Research, 256*, 353–358. https://doi.org/10.1016/j.psychres.2017.06.082

**How to cite this article:** Dombrovski, A. Y., & Hallquist, M. N. (2022). Search for solutions, learning, simulation, and choice processes in suicidal behavior. *Wiley Interdisciplinary Reviews: Cognitive Science, 13*(1), e1561. [https://doi.org/10.1002/wcs.1561](https://doi.org/10.1002/wcs.1561)