

Alpha Drive Theory

A Transdiagnostic White Paper on Perceived Agency, Action Selection, and Relief-Based Learning

White Paper / Preprint Version 1.0

Not peer reviewed

Lesa Warren, MSW

Independent researcher, United States

Prepared for public preprint posting on PsyArXiv

Author Note

Lesa Warren, MSW, holds a B.A. in Psychology from Baylor University (1992) and an M.S.W. in Direct/Clinical Practice from the University of Arkansas at Little Rock (1998), where she graduated with a 4.0 GPA and received the Joan Carder Alumni Scholarship for Academic Excellence. She was a Licensed Masters Social Worker in Texas from 1998 to 2001, and a Licensed Clinical Social Worker in Texas (2001–2011) and Arkansas (2001–2008; 2017–2025). From 1998 to 2007 she practiced social work in medical, forensic, child-welfare, and private-practice settings, including five years (2002–2007) as founder and director of the Behavioral Health Center of Texarkana, P.C., a multidisciplinary outpatient behavioral health center. She also served as adjunct faculty at Texas A&M University–Texarkana and as field instructor for graduate counseling and social work programs at Texas A&M and Baylor University. Earlier professional writing appeared in National Association of Social Workers outlets, including a cover article in *NASW Intersections in Practice* (Warren, 2007). Following a transition from clinical practice in 2009, she has conducted independent theoretical research, during which Alpha Drive Theory was formalized. The author has no institutional affiliation and reports no conflicts of interest. Correspondence may be directed to lesa.warren@gmail.com. The forthcoming book *The Unseen Why: The Hidden Logic Behind Everything You Do* is in preparation. Additional materials are available at TheAlphaDrive.com.

This document presents the scholarly, public-facing architecture behind Alpha Drive Theory. It is intended as a concise theoretical white paper, not as a substitute for the forthcoming trade book or for peer-reviewed empirical validation. Additional public-facing materials, figures, updates, and book information are available at TheAlphaDrive.com

Abstract

Alpha Drive Theory proposes a transdiagnostic model of behavior centered on perceived agency: the registration that one's action has produced credible effect in reality. The Alpha Drive (AD) is defined as a constant, upstream pressure toward interaction with reality. "Upstream" means earlier in the behavioral sequence, not more important morally or biologically. AD is proposed as the initiating pressure that makes behavioral engagement possible; other drives and needs then shape the domain, target, and urgency of that engagement. The Alpha Drive System (ADS) is the regulatory architecture through which that pressure is expressed, constrained by body state, selected into action, assessed, reinforced, and updated over time. The model treats perceived agency as a binary regulatory condition: at the Assess Gate, either both required checks pass and the person enters the Relief Window, or they do not and another iteration of The Loop runs. The model treats Assess Gate passage as binary, while acknowledging that the signals feeding that assessment may be continuous, graded, or probabilistic. Emotion has two functions in this architecture: an emotional read, evaluated as positive or negative at the Assess Gate, and an emotional posture, the specific qualitative form of the emotion carried forward into the next iteration through the Body State Gate. The Engagement Ladder organizes selected actions by the energy the system must marshal to reach toward shared reality: Shutdown, Internal Processing, Survey, Accommodation, Force, and Mastery. The Alpha Drive Continuum frames those actions as a behavioral continuum everyone occupies, with pathology defined not by location on the continuum but by stuckness, narrowing, and loss of behavioral flexibility. Borrowed agency, automatic defensive selection, defense mechanisms, DSM patterns, and dyadic/group dynamics are described as extensions of the same architecture rather than separate systems. The theory is offered as a synthetic, hypothesis-generating framework integrating perceived control, sense of agency, stressor controllability, allostasis, reinforcement learning, interoception, defense theory, and transdiagnostic psychopathology. Its central claim is simple: humans keep reaching for perceived agency, and when ordinary routes narrow, the system finds whatever route remains.

Keywords: Alpha Drive; perceived agency; perceived control; sense of agency; allostasis; relief learning; defense mechanisms; transdiagnostic psychopathology; behavioral regulation

Contents

- 1. Purpose and status of this white paper
- 2. The problem Alpha Drive Theory addresses
- 3. Core claim and key terms
- 4. Perceived agency as the regulated condition
- 5. Emotion: emotional read and emotional posture
- 6. The Alpha Drive System Loop
- 7. The Body State Gate and calibration
- 8. The Engagement Ladder
- 9. The Alpha Drive Continuum
- 10. Borrowed agency
- 11. Automatic control: fight/flight/freeze and defense mechanisms
- 12. Psychopathology and DSM diagnoses
- 13. Dyads, groups, and shared controllability
- 14. Clinical implications
- 15. Research agenda and falsifiable predictions
- 16. Boundaries, limitations, and publication status

- References
- Acknowledgment of AI use

1. Purpose and status of this white paper

This white paper introduces Alpha Drive Theory as a formal, public-facing theoretical framework. It is designed to establish the architecture of the model in a citeable form while preserving the forthcoming book as the primary reader-facing expression of the theory. It does not attempt to reproduce the book's narrative arc, personal material, vignettes, chapter structure, or full diagnostic walkthrough. This document provides the Alpha Drive mechanism.

The theory emerged from more than a decade of clinical practice followed by two decades of theoretical development, synthesis, and research. Its beginning observation was clinical: people with very different diagnoses, histories, and surface behaviors repeatedly seemed to be struggling with some version of the same problem - a disruption in the felt sense that their actions could produce meaningful effect. The present theory formalizes that observation into a regulatory model.

The theory is not presented as empirically proven. It is presented as synthetic and testable. Many of its components are anchored in established literatures - perceived control, sense of agency, controllability, allostasis, reinforcement learning, interoception, defense mechanisms, attachment, and transdiagnostic models. The novel claim is the architecture: perceived agency as the regulated condition, AD as constant upstream pressure toward interaction, the ADS Loop as the mechanism, and the Engagement Ladder as an energy-ordered map of selected actions.

Throughout this paper, Alpha Drive Theory refers to the broader explanatory framework, while model refers to the proposed regulatory architecture within that framework, including the Loop, Body State Gate, Engagement Ladder, Relief Window, Relief-Based Learning, and Drift.

2. The problem Alpha Drive Theory addresses

Clinical and behavioral sciences contain many powerful partial explanations. Cognitive-behavioral models explain the role of thought and belief. Reinforcement learning explains why immediately relieving behaviors become habits even when they carry delayed costs (Sutton & Barto, 2018). Stressor-controllability and learned-helplessness research explain why controllable and uncontrollable stressors produce different downstream effects (Amat et al., 2005; Maier & Seligman, 2016; Limbachia et al., 2021; Meine et al., 2021). Sense-of-agency research explains how action and outcome are linked in experience, including affective and bodily modulation of agency (Moore & Fletcher, 2012; Gentsch & Synofzik, 2014; Christensen et al., 2019; Villa et al., 2022). Interoception and pain-attention research explain how bodily state can capture attention and constrain regulatory options (Craig, 2002; Eccleston & Crombez, 1999; Khalsa et al., 2018). Attachment theory explains how early and adult relational systems shape later regulation (Ainsworth et al., 1978; Mikulincer & Shaver, 2016; Morris et al., 2007). Transdiagnostic, dimensional, and network models explain why DSM categories overlap and why symptoms cluster across diagnostic boundaries (Borsboom, 2017; Kotov et al., 2017). Perceptual control theory and free-energy/active-inference accounts provide neighboring formal vocabularies for feedback, perception, and action (Powers, 1973; Friston, 2010).

The problem is not that these models are wrong. The problem is that they often operate as parallel vocabularies. Clinicians and patients still need a single usable architecture for the moment-to-moment sequence: something becomes salient, body state constrains what can be selected, a target is selected; an action is selected and executed, the system evaluates whether it counted, relief may occur, and learning updates future selection.

Alpha Drive Theory proposes that the missing organizing variable is perceived agency. The claim is not merely that control matters; that is already well supported. The claim is that the need to register credible effect organizes action selection across domains, and that many behaviors that appear irrational, defensive, pathological, or self-defeating become mechanically coherent when understood as attempts to produce, protect, simulate, route, or restore perceived agency.

3. Core claim and key terms

Three terms must be separated because each does different work in the theory:

- **Alpha Drive (AD).** The constant upstream pressure toward interaction with reality. AD pushes. It does not evaluate, prefer, select, learn, or solve problems.
- **Alpha Drive System (ADS).** The regulatory architecture the AD pressure runs through. ADS integrates, selects, assesses, reinforces, and updates.
- **Alpha Drive Theory.** The theoretical claim that AD and ADS, structured as described, account for the initiation and regulation of behavior through perceived agency.

The Alpha Drive is proposed as first in a specific architectural sense. Human beings have many downstream drives and needs - hunger, attachment, safety, curiosity, sexuality, status, competence, belonging, and meaning. Alpha Drive Theory does not deny those systems or attempt to replace their literatures. It proposes that without an initiating pressure toward engagement with reality, none of those downstream drives has behavioral substrate through which to express. AD is the first domino. The downstream systems enter the ADS through the Body State Gate as state conditions, agency domains, salience patterns, and action tendencies.

AD pressure is constant. It does not rise and fall. What changes is whether perceived agency currently outcompetes that pressure in subjective experience. When perceived agency is registering, the pressure fades from attention because positive emotion and credible agency dominate. When perceived agency stops registering, the same constant pressure becomes salient, uncomfortable, and action-demanding. This distinction matters: intervention is not aimed at reducing AD. Intervention is aimed at creating credible, lower-cost routes to perceived agency.

The system is not solving life problems in the ordinary sense. It may appear to be solving a job problem, a relationship problem, a health problem, a parenting problem, or an identity problem. Mechanically, it is reaching for perceived agency under current Body State Gate conditions. Sometimes solving the life problem and restoring perceived agency align. Often they diverge. Many clinical problems live in that divergence.

4. Perceived agency as the regulated condition

Perceived agency (PA) is the system's registration that the organism's action has produced credible effect on reality. Reality may be external, interpersonal, bodily, future-oriented, symbolic, or internally modeled. An action can count because a person moved an object, influenced another person, changed a bodily state, shifted an internal model, recruited an intermediary, or meaningfully altered how the situation is represented.

PA is broader than self-efficacy, autonomy, locus of control, perceived control, and sense of agency, while overlapping with all of them (Bandura, 1977; Rotter, 1966; Gallagher et al., 2014; Moore & Fletcher, 2012; Wang et al., 2021). It includes the felt registration that I did that, and what I did made a difference in my reality. It is not a stable trait. It is a regulatory condition continuously sought, protected, simulated, or restored through action.

PA is binary in this model. It registers, or it does not. There is no partial PA and no partial opening of the Assess Gate. Variability appears elsewhere: in what the Body State Gate calibrates as enough under current conditions, in how long the person remains in the Relief Window once the Assess Gate opens, in the emotional posture carried forward, and in how Relief-Based Learning and Drift update future selection.

PA is structurally required where AD is operating. A person capable of experiencing agency constraint as negative and responding to it in some form cannot remain organized without some route to PA. If shared reality no longer produces PA through ordinary action, the system may produce it through stopping, internal simulation, repeated checking, self-adjustment, pressure, borrowed agency, substance use, altered relationship to reality, or countless other constrained routes. Some of those routes are costly. Some are catastrophic. They are still PA routes.

In this model, perceived agency is the only psychological requirement the system cannot simply override. Other drives and needs can be delayed, suppressed, redirected, or sacrificed under pressure. Perceived agency cannot be abandoned; if ordinary routes fail, the system must find some other route — internal, relational, borrowed, symbolic, defensive, or extreme — through which action can still register as having an effect.

5. Emotion: emotional read and emotional posture

Emotion is treated as the felt expression of perceived agency within the ADS. This is a narrow functional claim. It does not require a comprehensive theory of affect, and it does not settle debates between discrete-emotion theories, constructionist theories, appraisal theories, primary-process affect systems, or distinctions among mood, emotion, feeling, temperament, and personality (Ekman, 1992; Lazarus, 1991; Panksepp, 1998; Barrett, 2017). It specifies what emotion does in this architecture.

The model separates two emotion functions. The emotional read is binary: positive or negative. At the Assess Gate, a positive emotional read is necessary but not sufficient for PA to register. The credibility check must also register that the positive shift is attributable to the organism's own action producing effect. A negative emotional read signals that PA has not registered, has flipped to no, or is threatened with doing so.

This separation is compatible with sense-of-agency research showing that affective state, outcome valence, and non-motoric bodily cues can modulate the experience of agency rather than merely appearing after agency has already been registered (Christensen et al., 2019; Gentsch & Synofzik, 2014; Villa et al., 2022).

Emotional posture is description-rich and carried forward. Anger, fear, shame, grief, pride, satisfaction, awe, calm, disgust, contempt, relief, and joy are not merely colors added after the fact. They carry information about the domain in which perceived agency was gained, threatened, obstructed, restored, exceeded, or lost. The emotional posture coming out of one iteration becomes one of the primary inputs the Body State Gate reads when presenting the reachable action field for the next iteration.

This carry-forward function matters. Emotion does not merely report whether the last action counted; its specific posture helps bias what becomes reachable next. Anger tends to bias toward upper-rung action because the situation is being read as resistance. Shame tends to bias toward withdrawal, self-adjustment, or concealment. Fear biases toward defensive or safety-oriented action depending on escapability. Pride and satisfaction bias toward repetition or continued Mastery. Joy biases toward seeking more of what produced the unexpected positive resolution. Emotional posture is therefore a bridge from one Loop iteration to the next.

Positive and negative postures both have clinical information value.

- **Negative postures.** Fear points toward anticipated loss of control or danger; anger toward obstruction; shame toward threatened social standing or failed self-attribution; grief toward irreversible loss of an agency-producing bond, future, role, or world; helplessness toward absence of reachable routes; disgust toward bodily, moral, or relational threat; contempt toward degradation of the target.
- **Positive postures.** Pride indicates PA was achieved through demonstrated competence; satisfaction indicates action matching intent; warmth and connection indicate relational PA; relief indicates threat removal; calm or contentment indicates ordinary conditions registering as enough; joy or delight indicates PA exceeding expectation; awe indicates PA organized around something larger than the self.

The clinical implication is direct: what kind of positive emotional posture settles a person matters as much as what kind of negative emotional posture destabilizes them. Treatment is more precise when it asks not only what hurts, but what sort of positive emotional posture actually opens and sustains the Relief Window.

6. The Alpha Drive System Loop

The ADS Loop is an 11-step functional architecture. The steps are not proposed as fixed neural modules or discrete micro-states. They are a decomposition of a regulatory process that can run slowly with conscious participation or quickly below awareness. In the book, this is The Loop. In research language, it is a repeatable feedback cycle organized around perceived agency.

Salience Capture is used here as a functional description, not as a one-to-one claim about a specific neural network. Large-scale network accounts nevertheless provide plausible anchors for salience detection, network switching, executive regulation, and internal simulation dynamics (Seeley et al., 2007; Menon, 2011).

Step	Function
1. AD Trigger	Something happens – inside or outside the person – that disturbs the current sense of perceived agency. PA flips to ‘no’ or becomes threatened, AD pressure becomes salient, and the recurring Loop begins another iteration.
2. Salience Capture	Attention is allocated to an external event, internal signal, memory, threat, loss, need, or felt disturbance.
3. Body State Gate	Body state, energy economy, history, emotional posture, biochemistry, and current context are integrated into a reachable action field and calibration.
4. Identify Options	The system surveys what is reachable from the field the Body State Gate has presented.
5. Target Selection	The system selects what it will act on: body, self, other, situation, future, feeling, memory, story, or intermediary.
6. Action Selection	The system selects a specific behavior from the reachable field. Selection may be conscious, subconscious, or unconscious.

Step	Function
7. Action	The behavior is executed, outwardly or internally, slowly or faster than awareness can track.
8. Assess Gate	Two checks run simultaneously: the affective check and the credibility check. Both must pass.
9. Relief Window	When both checks pass, the person enters a state of relief from AD pressure. The window lasts only while conditions for staying hold.
10. Relief-Based Learning	The action pattern that enabled entry into the Relief Window is marked as effective and becomes more likely under similar conditions.
11. Drift	Baseline calibration and rung tendency update over repeated patterns, shaping what will count and what will be selected in the future.

The Assess Gate is the hinge. It runs two simultaneous checks. The affective check asks: did the action shift the emotional read positive? The credibility check asks: did this action produce perceived agency - the felt sense that I did that, and what I did made a difference in my reality? Both must pass. If either fails, the Assess Gate does not open and another iteration runs.

The Relief Window is not a gate. It is the state entered after the Assess Gate opens. The person remains in the Relief Window as long as conditions for staying hold: the action continues, no new disturbance captures salience, the original disturbance does not reassert, and body state continues to support the calibration. When conditions stop holding, the person leaves the Relief Window and another iteration begins.

Relief-Based Learning and Drift are distinct. Relief-Based Learning heavily weights what just happened: which action, in which domain, under which posture, produced or failed to produce a gate-opening. Drift accumulates over time, updating baseline calibration and rung tendency only when enough pattern has built up. Together they explain why patterns can form quickly in the moment and then consolidate slowly across months or years.

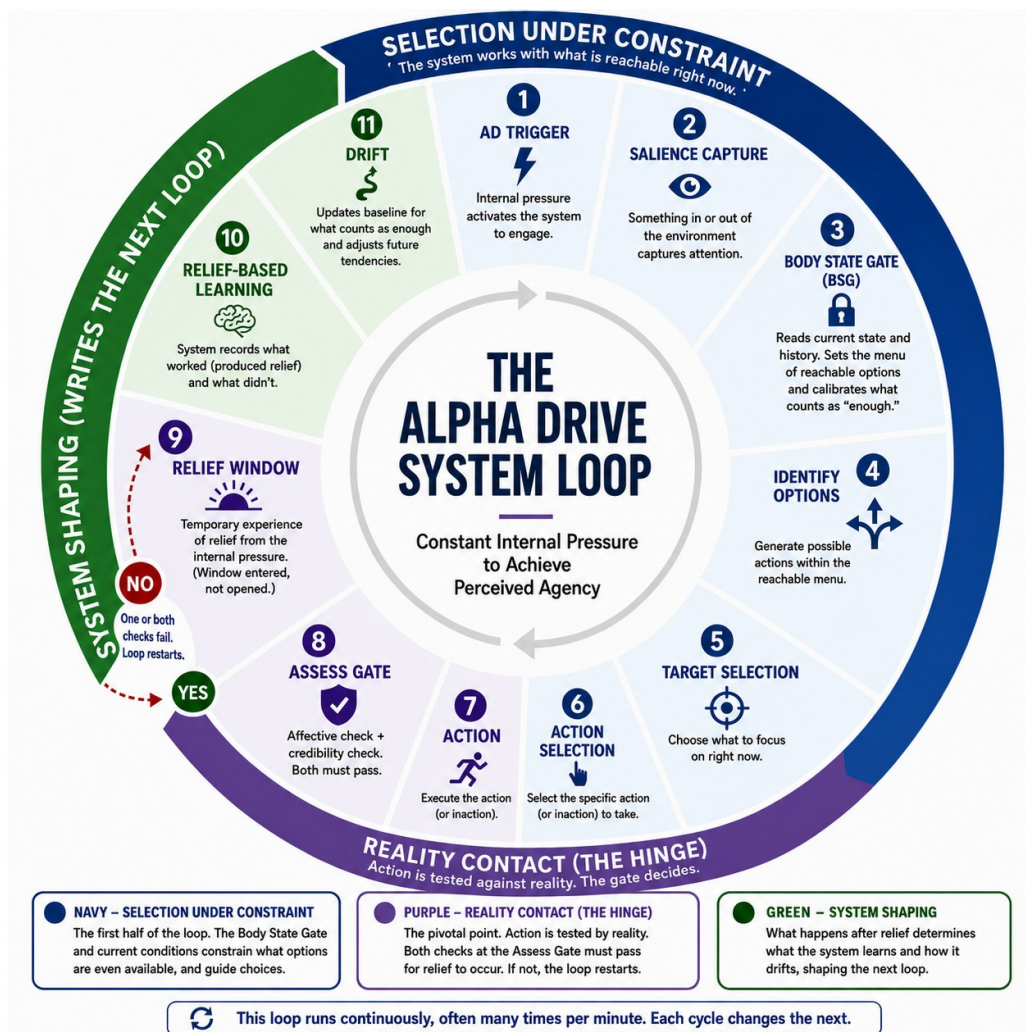


Figure 1. The Alpha Drive System Loop. Step 8, the Assess Gate, requires both the affective check and the credibility check to pass before the person enters the Relief Window. Adapted from the Alpha Drive Theory visual framework for white paper use.

7. The Body State Gate and calibration

The Body State Gate (BSG) is Step 3 of the Loop and one of the theory's central clinical mechanisms. The BSG reads what the organism currently is: sleep, pain, hunger, fatigue, inflammation, grief, panic, chronic stress, medication, substances, withdrawal, hormones, neurodevelopmental and neurocognitive conditions, emotional posture, history, culture, power dynamics, learned defaults, and current context. It does not moralize. It reads conditions.

Pain, fatigue, interoceptive signals, and other bodily conditions are especially important BSG inputs because they can capture attention and alter the reachable action field before conscious deliberation begins (Craig, 2002; Eccleston & Crombez, 1999; Khalsa et al., 2018).

The BSG produces two outputs. First, it presents the reachable action field: the menu of behaviors, targets, and actions actually available under current conditions. Second, it calibrates what the Assess Gate will recognize as enough for both checks to pass. Calibration means what the system has currently set as sufficient. The term threshold is avoided because it implies a single quantitative bar. Calibration is broader and more state-sensitive.

The same person facing the same problem can therefore be a different regulatory system on two different days. Knowledge may be the same. Values may be the same. The outside problem may be the same. But if pain is higher, sleep is lower, grief is acute, blood sugar is unstable, threat is present, or medication has shifted the biochemical field, the BSG presents a different menu and a different calibration. This is why advice like ‘just calm down’, ‘just think rationally’, or ‘just do what worked before’ often fails. The advice is aimed at the conscious aware self, but the action field is being generated upstream by the BSG.

Allostasis provides a useful physiological analogy. The body does not merely maintain fixed setpoints; it adjusts regulatory targets to meet current and anticipated demand, and repeated adjustment under chronic demand produces allostatic load (McEwen, 1998; Sterling, 2012). Alpha Drive Theory applies a parallel logic to behavioral regulation. Psychopathology is not the ADS breaking. It is the system adapting to chronic conditions, often skillfully, with costs that accumulate over time.

8. The Engagement Ladder

The Loop describes how behavior happens. The Engagement Ladder describes the selected actions themselves. It is an energy-ordered catalogue of action categories arranged by the amount of energy the system must marshal to reach toward shared reality in a way that can pass the Assess Gate. In this context, energy means mental-regulatory energy: the system’s available capacity to engage shared reality under current representation, meaning, and body-state conditions. Energy does not mean willpower, biological activation, sympathetic arousal, visible movement, or metabolic fuel alone. It means the system’s marshaled capacity to reach: to leave stopping, move beyond simulation, tolerate information, adjust the self, apply pressure, or meet shared reality directly. The selected action is what the system does with that energy; the energy is what makes that kind of action reachable.

The ladder is not a sequence, severity scale, moral ranking, or diagnostic filing cabinet. A healthy system can move across the entire ladder as conditions require. Stuckness is repertoire narrowing: the system has consolidated around too small a slice of one or more rungs, while other rungs have become less reachable or less practiced.

Rung	Strategy	Buffer	Energy and action description
1. Shutdown	Stopping	Stopping standing in for action	Minimal energy marshaled toward shared reality. The selected action is stopping: stillness, withdrawal, numbing, sleeping, collapsing, conserving. Shutdown is maximum constraint-compatible PA production, not system failure.

Rung	Strategy	Buffer	Energy and action description
2. Internal Processing	Thought	Thought standing in for action	Low energy marshaled toward shared reality; reach stays in the internal model. Selected actions include rumination, rehearsal, planning, fantasy, meaning-making, remembering, scrolling, or acquisition behaviors where PA comes from the simulation rather than external effect.
3. Survey	Information	Information standing in for action	Low-to-moderate energy marshaled; reach extends toward information about shared reality rather than direct change. Selected actions include checking, researching, monitoring, observing, reassurance seeking, due diligence, scanning, and asking around.
4. Accommodation	Self-adjustment	The self standing in for the situation	Moderate energy marshaled; reach moves through self-adjustment. Selected actions include yielding, appeasing, adapting, complying, masking, softening, compromising, or adjusting around a person or condition.
5. Force	Pressure	Pressure standing in for engagement	High energy marshaled; reach becomes pressure against resistance. Selected actions include confronting, demanding, coercing, challenging, setting hard limits, arguing, escalating, or applying pressure.
6. Mastery	Direct engagement	No rung-level buffer	Highest energy marshaled; reach meets shared reality directly. Selected actions include cooking, walking, relating, building, practicing, parenting, writing, creating, learning, repairing, being present, or skilled performance. Skill is downstream; the rung is direct engagement itself.

Each rung except Mastery uses a buffer. Shutdown uses stopping. Internal Processing uses thought. Survey uses information. Accommodation uses the self. Force uses pressure. Mastery has no rung-level buffer; the doing is the engagement. This buffer logic explains the ladder's energy ordering: as the buffer decreases, the system must marshal more energy toward shared reality contact.

Every rung is necessary. A system that cannot stop burns out. A system that cannot think ahead behaves recklessly. A system that cannot survey misses critical information. A system that cannot accommodate destroys relationships and groups. A system that cannot use Force is vulnerable when pressure or boundary-setting is necessary. A system that cannot access Mastery loses direct life contact. Health is not living at the top. Health is flexibility across the full range.

The rung does not determine morality. Force can be calibrated assertion or violence. Accommodation can be wisdom or self-erasure. Shutdown can be necessary rest or severe

collapse. Mastery can be life-giving direct engagement or a narrow excellence trap that leaves the rest of life uncalibrated. The rung tells us the action category. The outcome, cost, context, and downstream effects tell us whether the action fits.



Figure 2. The Engagement Ladder. The ladder is not a severity scale, moral hierarchy, or health ranking; health is flexibility across rungs.

9. The Alpha Drive Continuum

The Alpha Drive Continuum is the behavioral continuum created by The Loop and the Engagement Ladder. It is not a diagnostic scale, a severity line, or a moral ranking. Everyone whose Loop is running is on the Continuum all the time, because every selected action - including inaction - falls somewhere along the energy-and-engagement gradient. A person can move from Shutdown to Survey to Force to Mastery and back again within the same hour. The AD Continuum explains what the system is doing at a given point in time and why, not whether the person is pathological.

The point of the Continuum is movement. A healthy system is not a system that lives at Mastery or avoids the lower rungs. A healthy system can stop when stopping fits, think when thought fits, survey when information is needed, accommodate when adjustment is wise, use Force when pressure or boundary-setting is necessary, and engage directly when Mastery is reachable. The Continuum therefore describes behavioral flexibility: the breadth of the person's available repertoire across changing Body State Gate conditions.

Pathology enters when the system becomes stuck. Stuckness means the reachable action field has narrowed so far that the system repeatedly selects from a small slice of the Continuum even when other action categories would fit the situation better. The problem is not the rung itself. Shutdown is not pathological; living with Shutdown as the only reliable route to PA can become pathological. Force is not pathological; Force as the only reliable route to PA can become destructive. Mastery is not pathology-proof; a narrow slice of Mastery can become the only place the person can register PA while the rest of life goes uncalibrated.

This distinction matters for diagnosis. DSM labels can be read as descriptions of recurring patterns on the AD Continuum - position-based patterns, looping patterns, oscillating patterns, and patterns that alter the Loop itself - but the Continuum is not a new diagnostic filing system. It is a process map. At any moment, it helps ask and answer: Which action category is the system selecting? What did the Body State Gate make reachable? What emotional posture is being carried forward? What did the Assess Gate recognize as enough? What is Relief-Based Learning consolidating, and where is Drift pulling the baseline?

The clinical goal is not to push everyone upward. The goal is restored flexibility: a wider menu of available action strategies, enough body-state capacity to reach them, and enough repeated gate-openings across the Continuum that the system no longer has to keep returning to one costly route. In Alpha Drive terms, health is not a location. Health is the ability to move.

10. Borrowed agency

Borrowed agency is a way of running the Loop, not a rung. In borrowed agency, the person produces PA by routing action through another person, group, institution, role, deity, ideology, legal structure, treatment system, expert, or other intermediary. The selected

action may fall on any rung, but cost-bearing capacity or effectiveness is imported through the attachment to something outside the individual system.

Borrowed agency has two functions. First, it imports capacity: a depleted person can access action that would not be reachable alone because another person, group, or structure supports it, lowers the cost, or carries part of the reach. Second, it routes action through an intermediary: the lawyer sends the letter, the doctor prescribes the treatment, the church carries the bereaved person, the group supplies certainty, the spouse handles the decision. The credibility check can still pass because the person's system authored the routing.

Borrowed agency is developmentally normal. Infants and children depend on caregivers to co-regulate distress, scaffold effective action, and provide a secure base from which exploration becomes possible; adult attachment systems can continue to shape regulation, threat response, and relational agency across the lifespan (Ainsworth et al., 1978; Mikulincer & Shaver, 2016; Morris et al., 2007). In healthy development, borrowed agency helps the child's own ladder come online. In constrained development or later trauma, borrowed agency can become the only reliable route to PA.

The distinction between functional and stuck borrowed agency is whether the attachment feeds the person energy back into their own repertoire or replaces that repertoire. A support group that helps a bereaved person eat, sleep, work, connect, and act again is functional borrowed agency. A group, leader, partner, ideology, or institution that becomes the only route through which PA can register has become structurally load-bearing and potentially costly. This mechanism helps explain dependency, abusive attachment patterns, high-control groups, ideological fusion, radicalization, and some forms of institutional loyalty without reducing those phenomena to moral weakness or irrationality.

11. Automatic control: fight/flight/freeze and defense mechanisms

Behavioral selection may be conscious, subconscious, or unconscious. Conscious deliberation is one form of selection, not the prototype of selection. Most behavior, most of the time, is selected below full awareness from a reachable field already shaped by the Body State Gate. This is not moral voluntariness. It is functional nonrandomness: the system selects from what is reachable, not from all actions theoretically possible.

Fight, flight, and freeze are automatic responses produced under acute threat conditions. They do not bypass The Loop. Rather, the Body State Gate narrows the reachable field so tightly that option identification, target selection, and action selection run faster than awareness can track. Fight usually lands on Force. Freeze lands on Shutdown. Flight lands on escape action, which may involve Accommodation, Force, or other rungs depending on context. Defensive-response and threat-imminence literatures provide compatible anchors for this framing (LeDoux & Pine, 2016; Mobbs et al., 2007; Roelofs, 2017; Coimbra et al., 2023).

Psychodynamic defense mechanisms belong in the same architecture. Repression, projection, rationalization, denial, displacement, intellectualization, dissociation, self-blame, premature forgiveness, defensive aggression, and similar processes are learned automatic or subconscious action patterns that shift the system when direct engagement with internal or external content would threaten PA. They are not signs that the system is broken. They are evidence that the system learned a route to PA under constraint.

Defense mechanisms do not feel like strategies. They feel like reality. By the time the conscious, aware self notices the action, the system may already have altered attention, meaning, self-attribution, emotional posture, and target selection. This explains the

subjective discontinuity of 'I snapped', 'I went blank', 'that was not me', or 'I heard myself say it before I decided'. The conscious, aware self encounters an action already in progress because the reachable menu was narrow and selection was fast.

Insight helps, but insight is not sufficient. Insight can name the pattern, but it does not automatically change body state, widen the reachable field, or make a different action pass the Assess Gate under pressure. Change requires new experience that opens the Assess Gate under real conditions through a different route.

12. Psychopathology and DSM diagnoses

Alpha Drive Theory preserves DSM diagnoses as clinically useful descriptors while reframing them as patterns of constrained action selection, Relief-Based Learning, and Drift. DSM categories remain important for communication, service access, risk recognition, research grouping, and shared clinical language. The theory does not replace DSM. It offers a process layer underneath it.

This process framing is compatible with, but not identical to, dimensional and network approaches to psychopathology, which also emphasize comorbidity, symptom covariation, and the limits of traditional categorical diagnosis (Borsboom, 2017; Kotov et al., 2017).

The Alpha Drive Continuum described above is the behavioral map on which DSM patterns can be placed without turning the continuum into pathology. Diagnosis enters when movement narrows: when the reachable action field becomes rigid, costly, impairing, dangerous, or clinically significant. Psychopathology is therefore not entry onto the Continuum; it is loss of flexibility on a continuum everyone already occupies.

Anxiety patterns often involve high salience, constrained cost tolerance, and repeated Survey or Internal Processing. Compulsive patterns often involve repeated action in which Relief Windows are brief or fail to sustain. Depressive patterns often involve narrowed cost tolerance in which Shutdown or Internal Processing become the most reachable routes to PA. Trauma patterns often involve Body State Gate conditions organized around historical threat data. Mania involves altered body-state conditions that make higher-cost action categories feel reachable at unusual speed and breadth. Substance use can create long or intense Relief Windows under altered body-state conditions, powerfully reinforcing the substance-linked route when non-substance routes are unavailable.

Psychosis, dissociation, dementia, autism, ADHD, intellectual disability, brain injury, and other neurodevelopmental or neurocognitive conditions require care in this framing. The model applies to behavioral regulation, not as a replacement for biology. The Body State Gate and biochemistry are not separable. Neurological and developmental conditions can change what the Body State Gate has to work with: sensory load, memory, inhibition, salience, shared-reality reach energy, communication, prediction, and action availability. Those changes may make PA routes harder to see from the outside, but they do not imply an absent or broken Alpha Drive System. Wherever a person remains capable of interacting with reality - external, relational, bodily, or internally modeled - the system should be presumed to be reaching for perceived agency through whatever routes remain available. Resistance to restraint, unwanted handling, forced compliance, environmental overwhelm, or removal of a preferred route should therefore be read as agency-protective behavior, not as meaningless dysregulation. The clinical question is not whether the person has a working ADS; it is what route to PA remains reachable under that person's current neurological and body-state conditions.

13. Dyads, groups, and shared controllability

The individual architecture scales because human systems interact. When two or more people engage, their Loops don't merge. Each system runs its own Loop. The output of each system's Action step becomes input into the other systems at AD Trigger or Salience Capture. This Cross-System Input Rule is the foundation of dyadic and group analysis.

A stable dyad is one in which both systems can pass the Assess Gate in the same shared space. This is shared controllability: both people can register that their actions matter without one person's PA requiring the other person's loss of PA. Shared controllability is not harmony, agreement, similarity, or love. Love can motivate effort and keep people present, but love is not a control mechanism. The Loop runs on whether the Assess Gate passes.

Different strategies can fit together without creating conflict. One person may lead while another adjusts; one may gather information while another acts; one may push while another slows the pace. The problem begins when one system's PA-restoring action repeatedly becomes PA-reducing input for the other system. At that point, shared controllability fails. Each action may still make sense inside its own Loop, but the interaction becomes destabilizing because one system can settle only by unsettling the other.

Repeated dyadic interactions produce Cross-System Drift. Each system's Drift step integrates what the other system's outputs have taught it. The next argument starts faster, escalates faster, or collapses faster not because either person chose that pattern, but because both systems updated from prior interactions. Asymmetric cost compounds this: the same interaction does not cost both bodies the same amount. Over time the pattern tends to organize around the actions that one system can produce and sustain most easily, even when those actions are costly for the other system, unless the dyad actively creates boundaries around that asymmetry.

In larger group settings, borrowed agency can contribute to group cohesion by coordinating the systems of multiple individuals around a shared person, institution, ideology, purpose, or identity. Once identification with the group is established, interactions between the group and outsiders can operate much like dyadic interactions: the group's outputs become inputs into other systems, and other systems' outputs become inputs into the group. The Cross-System Input Rule and Cross-System Drift can therefore apply to the group as a functional unit, while shared controllability remains the condition for stability. In this sense, Alpha Drive Theory can scale from individuals to dyads, groups, and institutions without changing the basic architecture.

14. Clinical implications

The primary clinical implication is a shift from surface form to Loop function. A behavior may be frightening, socially unacceptable, dangerous, costly, or catastrophic while still being the system's currently reachable route to PA. This does not excuse harm. It explains why behavior that makes no sense from outside can be powerfully reinforced inside the system.

An ADS-informed assessment asks:

- What disturbed the person's ability to remain in the Relief Window?
- What emotional read and emotional posture appeared?
- What did the Body State Gate have to integrate - sleep, pain, grief, hunger, threat, attachment, power dynamics, history, medication, substance use, or current cost tolerance?
- What action field was reachable?
- What target was selected?

- What action was selected?
- Which rung did the action fall on?
- Did the affective check and credibility check both pass?
- What enabled entry into the Relief Window?
- What did Relief-Based Learning consolidate?
- What is Drift likely building?

Treatment can enter at many points. It can improve body-state conditions, reduce pain or sleep debt, add structure, widen the reachable action field, use borrowed agency temporarily, increase conscious deliberation before action selection, change the target, build repertoire across the Engagement Ladder, reduce reinforcement of costly actions, and create small PA-producing experiences that can be repeated until Drift shifts.

The model also clarifies why removing a symptom without replacing its PA function often produces substitution. If checking, substance use, self-harm, rumination, withdrawal, aggression, or violence is the only reliable gate-opening action currently reachable, removing the behavior does not remove AD pressure. The system will select the next most available route. Intervention must build a route that can actually pass both checks under the person's real Body State Gate conditions without high cost.

The model also clarifies a known clinical danger: increasing a person's capacity to reach toward shared reality before a credible PA pathway exists can destabilize the system. Evolutionary accounts of low mood and depressive withdrawal have proposed that depressive states may, in some contexts, limit futile or dangerous engagement, conserve effort, reduce exposure to further social or environmental cost, or disengage the person from unreachable goals (Nesse, 2000; Andrews & Thomson, 2009; Durisko et al., 2015). ADS makes a narrower mechanical claim: when Shutdown or Internal Processing is the only route through which PA is currently registering, sudden increases in energy - through antidepressant activation, stimulant effects, substances, mania, or external pressure - can expand reach before the person has a safe strategy for making that reach count. The person can now act farther into shared reality while still experiencing reality as nonresponsive. This is why treatment must pair activation with immediate, credible, lower-cost PA pathways rather than simply increasing activity or energy (Hammad et al., 2006; Luft et al., 2018).

Suicidal ideation also requires distinction from suicidal action. Ideation may function as an Internal Processing action: the thought that suicide exists as an available option can itself produce PA when all other routes feel blocked. This can temporarily reduce panic or helplessness without meaning suicidal action is inevitable. Escalation risk rises when ideation becomes the only reliable PA route, when lower-cost alternatives remain unavailable, and especially when energy or reach increases enough for action to become selectable. Clinical intervention therefore should not treat ideation only as a prediction of action; it should ask what PA function the ideation is serving and build safer, lower-cost routes that can pass the Assess Gate before the system escalates.

This functional framing does not replace standard suicide-risk assessment, safety planning, or urgent intervention when imminent risk is present.

Non-suicidal self-harm should be treated as a distinct mechanism. In ADS terms, self-harm can externalize an intolerable internal state into something concrete, controllable, and attributable: I caused this sensation, I can see it, and I can stop it. The behavior may produce a brief Relief Window because the person has converted diffuse emotional pain, numbness, or internal chaos into a controlled action with visible effect. That does not make

the behavior healthy or safe. It explains why it can be powerfully reinforced and why removing it without replacing the PA function often fails.

The behavior may be costly, frightening, or catastrophic, but the system is not selecting cost for its own sake. It is preserving the non-negotiable requirement that some action, somewhere, still register as mattering. Treatment therefore cannot only remove the behavior. It has to create a safer, more flexible route through which perceived agency can register.

15. Research agenda and falsifiable predictions

Alpha Drive Theory is designed to generate testable predictions. The constructs require operationalization, but many neighboring measures already exist: perceived-control scales, agency tasks, intentional binding paradigms, controllability manipulations, ecological momentary assessment, physiological stress markers, interoceptive measures, sleep/pain indices, behavioral flexibility tasks, and clinical coding of action categories.

- **P1. Strategy shift.** Holding objective outcomes constant, experimentally reducing perceived controllability should shift selection toward narrower, more defensive, or more rapidly selected actions and away from exploratory or flexible actions.
- **P2. Action-field narrowing.** Under low controllability, high threat, pain, sleep loss, grief, intoxication, withdrawal, or other Body State Gate constraints, participants should report and display fewer reachable actions and greater reliance on previously reinforced rungs.
- **P3. Emotional posture carry-forward.** The emotional posture following one action should predict the reachable action field in the next iteration. Anger should bias upper-rung action; shame should bias withdrawal or self-adjustment; fear should bias defensive or safety-oriented action depending on escapability; pride and satisfaction should bias repetition or Mastery-related continuation.
- **P4. Relief-Based Learning.** Actions that enable entry into the Relief Window should show increased reuse probability under similar conditions, even when delayed consequences are negative. The effect should be stronger when the Relief Window is sustained longer.
- **P5. Target substitution.** When one target does not produce PA, the system should reach for other targets that can. Substitution may involve a different target in the same domain, a different domain, a different rung, borrowed agency, or repeated attempts on the same target when repertoire is narrow.
- **P6. Drift.** Repeated gate-opening on a rung should increase future selection of actions on that rung. Repeated success on high-cost shared-reality engagement should shift calibration upward; repeated low-cost gate-openings should shift calibration downward; repeated failure should narrow expected agency.
- **P7. Treatment response.** Interventions that create small, reachable PA pathways that pass both checks under actual Body State Gate conditions should reduce symptom substitution more reliably than interventions that suppress surface behavior without replacing the PA function.
- **P8. Defense mechanism prediction.** Defense mechanisms should predict sudden rung transitions under conditions that narrow the reachable action field. Subjective reports should often include discontinuity: 'I snapped', 'I went blank', or 'I was already doing it before I realized.'
- **P9. Borrowed-agency prediction.** Under depletion or high constraint, people should show increased reliance on intermediaries that lower shared-reality reach cost or route

action. Borrowed agency should be functional when it restores direct repertoire and costly when it becomes the only PA route.

- **P10. Substance-use prediction.** Substance use should be especially reinforcing when it sustains entry into the Relief Window under conditions where non-substance routes are unavailable. As substance use damages other PA pathways, the system should increasingly return to the substance because it remains the most reliable gate-opening action pattern.

16. Boundaries, limitations, and publication status

This white paper proposes a theoretical architecture. It does not provide direct empirical validation of the full ADS Loop as a unit. Several constructs require operationalization, including PA registration, emotional read, emotional posture, Body State Gate, Relief Window duration, calibration, rung selection, borrowed agency, shared controllability, and Drift.

Relation to neighboring frameworks

Several neighboring frameworks already formalize parts of this territory. Perceptual control theory describes behavior as feedback regulation organized around the control of perception (Powers, 1973). Free-energy and active-inference accounts provide a broader computational vocabulary for perception and action as inference processes that reduce prediction error or free energy (Friston, 2010). Self-determination theory identifies autonomy, competence, and relatedness as central psychological needs (Deci & Ryan, 2000), and hierarchy-of-needs models provide an older motivational map of organized human needs (Maslow, 1943). Alpha Drive Theory does not claim to replace these frameworks. Its narrower proposal is that the Alpha Drive is upstream and perceived agency functions as the regulated condition around which action selection, relief, and Drift can be clinically organized.

A major limitation is that Alpha Drive Theory has not yet been formalized as a computational model. The ADS Loop is specified functionally, but it is not yet written in mathematical or executable form. In that respect, the theory remains less formalized than perceptual control theory and active-inference/free-energy approaches, which already include more explicit computational machinery (Powers, 1973; Friston, 2010). The present paper therefore offers a functional clinical architecture that still requires formal modeling, measurement development, and empirical testing.

The upstream claim is strong. Alpha Drive Theory argues that AD sits above other drives and needs in the architecture of behavioral initiation. That claim invites engagement from self-determination theory, attachment theory, primary-process affect theory, appraisal theory, constructionist emotion theory, hierarchy-of-needs models, and evolutionary accounts of mood and motivation. This paper does not attempt to reconcile every framework. It argues that their constructs can be interpreted as downstream domains, state conditions, action tendencies, developmental histories, or implementation-level descriptions inside a broader PA-regulation architecture.

The model also risks overgeneralization if treated as a total explanation for every detail of behavior or psychopathology. That is not its purpose. Alpha Drive Theory is a behavioral-regulation model. It must interact with physiology, social structure, culture, trauma history, neurodevelopment, medication effects, poverty, disability, oppression, and material

conditions. The theory is strongest when used to ask what the system is reaching for under real constraints, not when used as a one-size answer to every question.

Finally, the language of perceived agency can be misunderstood as voluntarism. To avoid that, the theory uses selection rather than choice. Selection includes conscious, subconscious, and unconscious processes. The theory does not claim people deliberately produce symptoms. It claims the system selects actions that have produced or protected PA under available conditions.

Relation to the forthcoming book

This white paper is a scholarly proof-of-concept and public timestamp for the architecture. The forthcoming book, *The Unseen Why: The Hidden Logic Behind Everything You Do*, is the primary reader-facing version of the theory. The book develops the model through narrative, vignettes, chapter-length treatment of each rung, relational dynamics, diagnostic chapters, treatment approaches, and practical self-recognition. This white paper intentionally does not reproduce that material.

References

- Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). Patterns of attachment: A psychological study of the Strange Situation. Erlbaum.
- Amat, J., Baratta, M. V., Paul, E., Bland, S. T., Watkins, L. R., & Maier, S. F. (2005). Medial prefrontal cortex determines how stressor controllability affects behavior and dorsal raphe nucleus. *Nature Neuroscience*, 8(3), 365–371. <https://doi.org/10.1038/nn1399>
- Andrews, P. W., & Thomson, J. A., Jr. (2009). The bright side of being blue: Depression as an adaptation for analyzing complex problems. *Psychological Review*, 116(3), 620–654. <https://doi.org/10.1037/a0016242>
- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.
- Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, 16(1), 5–13. <https://doi.org/10.1002/wps.20375>
- Christensen, J. F., Di Costa, S., Beck, B., & Haggard, P. (2019). I just lost it! Fear and anger reduce the sense of agency: A study using intentional binding. *Experimental Brain Research*, 237, 1205–1212. <https://doi.org/10.1007/s00221-018-5461-6>
- Coimbra, B. M., Hoeboer, C. M., van Zuiden, M., Williamson, R. E., D'Elia, A. T., Mello, A. F., Mello, M. F., & Olff, M. (2023). The relationship between tonic immobility and the development, severity, and course of posttraumatic stress disorder: Systematic and meta-analytic literature review. *Journal of Anxiety Disorders*, 97, 102730. <https://doi.org/10.1016/j.janxdis.2023.102730>
- Craig, A. D. (2002). How do you feel? Interoception: The sense of the physiological condition of the body. *Nature Reviews Neuroscience*, 3(8), 655–666. <https://doi.org/10.1038/nrn894>
- Deci, E. L., & Ryan, R. M. (2000). The what and why of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

- Durisko, Z., Mulsant, B. H., & Andrews, P. W. (2015). An adaptationist perspective on the etiology of depression. *Journal of Affective Disorders*, 172, 315–323. <https://doi.org/10.1016/j.jad.2014.09.032>
- Eccleston, C., & Crombez, G. (1999). Pain demands attention: A cognitive-affective model of the interruptive function of pain. *Psychological Bulletin*, 125(3), 356–366. <https://doi.org/10.1037/0033-2909.125.3.356>
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3–4), 169–200. <https://doi.org/10.1080/02699939208411068>
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, 11, 127–138. <https://doi.org/10.1038/nrn2787>
- Gallagher, M. W., Bentley, K. H., & Barlow, D. H. (2014). Perceived control and vulnerability to anxiety disorders: A meta-analytic review. *Cognitive Therapy and Research*, 38, 571–584. <https://doi.org/10.1007/s10608-014-9624-x>
- Gentsch, A., & Synofzik, M. (2014). Affective coding: The emotional dimension of agency. *Frontiers in Human Neuroscience*, 8, 608. <https://doi.org/10.3389/fnhum.2014.00608>
- Hammad, T. A., Laughren, T., & Racoosin, J. (2006). Suicidality in pediatric patients treated with antidepressant drugs. *Archives of General Psychiatry*, 63(3), 332–339. <https://doi.org/10.1001/archpsyc.63.3.332>
- Khalsa, S. S., Adolphs, R., Cameron, O. G., Critchley, H. D., Davenport, P. W., Feinstein, J. S., Feusner, J. D., Garfinkel, S. N., Lane, R. D., Mehling, W. E., Meuret, A. E., Nemeroff, C. B., Oppenheimer, S., Petzschner, F. H., Pollatos, O., Rhudy, J. L., Schramm, L. P., Simmons, W. K., Stein, M. B., ... Paulus, M. P. (2018). Interoception and mental health: A roadmap. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 3(6), 501–513. <https://doi.org/10.1016/j.bpsc.2017.12.004>
- Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., Brown, T. A., Carpenter, W. T., Caspi, A., Clark, L. A., Eaton, N. R., Forbes, M. K., Forbush, K. T., Goldberg, D., Hasin, D., Hyman, S. E., Ivanova, M. Y., Lynam, D. R., Markon, K., ... Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, 126(4), 454–477. <https://doi.org/10.1037/abn0000258>
- Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press.
- LeDoux, J. E., & Pine, D. S. (2016). Using neuroscience to help understand fear and anxiety: A two-system framework. *American Journal of Psychiatry*, 173(11), 1083–1093. <https://doi.org/10.1176/appi.ajp.2016.16030353>
- Limbachia, C., Morrow, K., Khibovska, A., Meyer, C., Padmala, S., & Pessoa, L. (2021). Controllability over stressor decreases responses in key threat-related brain areas. *Communications Biology*, 4, 42. <https://doi.org/10.1038/s42003-020-01537-5>
- Luft, M. J., Lamy, M., DelBello, M. P., McNamara, R. K., & Strawn, J. R. (2018). Antidepressant-induced activation in children and adolescents: Risk, recognition and management. *Current Problems in Pediatric and Adolescent Health Care*, 48(2), 50–62. <https://doi.org/10.1016/j.cppeds.2017.12.001>
- Maier, S. F., & Seligman, M. E. P. (2016). Learned helplessness at fifty: Insights from neuroscience. *Psychological Review*, 123(4), 349–367. <https://doi.org/10.1037/rev0000033>

- Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4), 370–396. <https://doi.org/10.1037/h0054346>
- McEwen, B. S. (1998). Stress, adaptation, and disease: Allostasis and allostatic load. *Annals of the New York Academy of Sciences*, 840, 33–44. <https://doi.org/10.1111/j.1749-6632.1998.tb09546.x>
- Meine, L. E., Meier, J., Meyer, B., & Wessa, M. (2021). Don't stress, it's under control: Neural correlates of stressor controllability in humans. *NeuroImage*, 245, 118701. <https://doi.org/10.1016/j.neuroimage.2021.118701>
- Menon, V. (2011). Large-scale brain networks and psychopathology: A unifying triple network model. *Trends in Cognitive Sciences*, 15(10), 483–506. <https://doi.org/10.1016/j.tics.2011.08.003>
- Mikulincer, M., & Shaver, P. R. (2016). *Attachment in adulthood: Structure, dynamics, and change* (2nd ed.). Guilford Press.
- Mobbs, D., Petrovic, P., Marchant, J. L., Hassabis, D., Weiskopf, N., Seymour, B., Dolan, R. J., & Frith, C. D. (2007). When fear is near: Threat imminence elicits prefrontal-periaqueductal gray shifts in humans. *Science*, 317(5841), 1079–1083. <https://doi.org/10.1126/science.1144298>
- Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: A review of cue integration approaches. *Consciousness and Cognition*, 21(1), 59–68. <https://doi.org/10.1016/j.concog.2011.08.010>
- Morris, A. S., Silk, J. S., Steinberg, L., Myers, S. S., & Robinson, L. R. (2007). The role of the family context in the development of emotion regulation. *Social Development*, 16(2), 361–388. <https://doi.org/10.1111/j.1467-9507.2007.00389.x>
- Nesse, R. M. (2000). Is depression an adaptation? *Archives of General Psychiatry*, 57(1), 14–20. <https://doi.org/10.1001/archpsyc.57.1.14>
- Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. Oxford University Press.
- Powers, W. T. (1973). *Behavior: The control of perception*. Aldine.
- Roelofs, K. (2017). Freeze for action: Neurobiological mechanisms in animal and human freezing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1718), 20160206. <https://doi.org/10.1098/rstb.2016.0206>
- Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied*, 80(1), 1–28. <https://doi.org/10.1037/h0092976>
- Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., Reiss, A. L., & Greicius, M. D. (2007). Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, 27(9), 2349–2356. <https://doi.org/10.1523/JNEUROSCI.5587-06.2007>
- Sterling, P. (2012). Allostasis: A model of predictive regulation. *Physiology & Behavior*, 106(1), 5–15. <https://doi.org/10.1016/j.physbeh.2011.06.004>
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.

- Villa, R., Ponsi, G., Scattolin, M., Panasiti, M. S., & Aglioti, S. M. (2022). Social, affective, and non-motoric bodily cues to the Sense of Agency: A systematic review of the experience of control. *Neuroscience & Biobehavioral Reviews*, 142, 104900.
<https://doi.org/10.1016/j.neubiorev.2022.104900>
- Wang, K. S., Yang, Y.-Y., & Delgado, M. R. (2021). How perception of control shapes decision making. *Current Opinion in Behavioral Sciences*, 41, 85–91.
<https://doi.org/10.1016/j.cobeha.2021.04.003>

Acknowledgment of AI use

During preparation of this white paper, the author used large language models (OpenAI ChatGPT models and Anthropic Claude models, accessed through their respective web interfaces) for writing support, including drafting passages, revising text, clarifying phrasing, organizing sections, comparing versions, and stress-testing theoretical claims for internal consistency. The theoretical framework presented here emerged from the author's eleven years of clinical practice and has been developed, tested, and refined over the two decades since; AI tools served as writing aids and editorial sounding boards, not as originators of theoretical content. The author reviewed and edited the text and takes full responsibility for the accuracy and integrity of the work. No AI tool was used to generate or analyze empirical data.