Commentary

Awareness, Automation, and the Illusion of Free Will: Rethinking Libet through Trilogy Theory

Ashkan Farhadi[1,2]

1. Digestive Disease Center, MemorialCare Medical Group, United States; 2. Department of Medicine, University of California, Irvine, United States

This article explores learning, behavioral conditioning, and the neuroscience of volition through the lens of the Trilogy Theory of Consciousness (TTC), a model that distinguishes between awareness-based decision-making and algorithm-driven automation. The essay draws parallels to Pavlovian conditioning and expands into a broader framework that redefines learning as a transition from Awareness-Based Choice Selection (ABCS) to Selection of Choice Based on Algorithm (SCBA). TTC outlines a three-stage decision-making process—preselection (integration of informational and emotional intelligence), selection (where awareness momentarily guides choice via ABCS), and post-selection consolidation (the encoding of behavior into internalized algorithms).

Incorporating a reinterpretation of Benjamin Libet's famous[1] experiment, the article argues that what was previously taken as evidence against free will may instead reflect a misidentification of SCBA as volitional choice. Libet's participants engaged in repetitive tasks that likely bypassed true awareness-based intention, leading to the false conclusion that unconscious neural activity precedes all decision-making. TTC clarifies how awareness initiates intention and remains active through feedback loops, offering an updated model of volition that reconciles neuroscience with subjective experience.

Finally, the article proposes experimental approaches—such as digital piano performance tasks and neuroimaging—to empirically observe the transition from conscious choice to algorithmic behavior. By reframing learning and conditioning as transformations of awareness into structured behavioral automation, TTC offers a new framework for understanding how the mind is wired—and how it can be consciously rewired for adaptability, growth, and agency.

 doi.org/10.32388/DLUQV1

A Forgotten Pond Story

I hadn't cleaned my backyard pond for quite some time—years, in fact. Nor had I fed the fish, assuming they were long gone. After all, neglect over a couple of years should have been enough to wipe out the tiny ecosystem I once nurtured. But when I finally decided to clean the pond, I was surprised to find that a few small mosquito fish had survived. Encouraged, I added a few new goldfish to the pond and, out of habit more than hope, resumed my old routine: feeding them every morning before heading to work.

In the past, the fish would eagerly gather near the surface at the first sign of me approaching, anticipating food. But now, things were different. The fish scattered at my presence, hiding under shrubs and avoiding the food entirely.

Still, I knew from experience—and a bit of faith—that this would change with persistence. And I was right. After about ten days, I noticed a subtle shift. One or two fish no longer fled when I neared the pond. Instead, they gathered where the food would drop and began nibbling even while I stood nearby. Each day, more fish joined them, until soon enough, they all swam toward the food the moment I approached—racing, competing, no longer seeing me as a threat, but as a signal.

This progression was strikingly familiar. It reminded me of Pavlov's classic conditioning experiments, in which dogs learned to associate the sound of a bell with food. But this wasn't just a replication of Pavlov—it was a small backyard mirror of a broader truth: learning and adaptation are foundational to both animal and human behavior.

What I witnessed in my backyard pond was more than just a behavioral shift in a group of fish — it was a living reenactment of one of the most well-known experiments in behavioral psychology. Over a century ago, Ivan Pavlov observed similar learning processes in dogs, laying the foundation for our modern understanding of **classical conditioning**. His work revealed that, through repetition and association, animals could learn to respond to previously neutral stimuli as if they were biologically significant. This was the first glimpse into how the **external world rewires the internal patterns of response**.

Classical Conditioning (Pavlov's Experiment)

In his famous experiment, Pavlov rang a bell every time he fed his dogs. Initially, the bell meant nothing to them. But over time, the dogs began to salivate at the sound of the bell alone — even when no food was presented. This anticipatory reaction, known as a **conditioned response**, emerged from the association between the neutral stimulus (bell) and the unconditioned stimulus (food). Pavlov's work was revolutionary not only because it explained animal learning, but because it offered a scientific model for how **habits and reflexes are formed through repetition**, bypassing volitional choice or modifying a complex reflexes.

Since Pavlov's time, countless studies have expanded our understanding of **learning, memory, and behavioral conditioning** in both animals and humans. A 2018 study published in *Nature Neuroscience* demonstrated how synaptic plasticity in the amygdala supports the formation of fear conditioning, linking emotional learning to specific neural circuits[2]. Another recent paper in *Neuron* (2021) identified how dopamine signals not only reinforce learning but also encode the **transition from deliberate to automatic behavior**, essentially capturing the neurobiological process underlying **behavioral automation**[3]. In humans, fMRI studies[4] show how regions like the prefrontal cortex are active during **initial choice-making** but become less involved once behaviors become habitual — again supporting the view that **awareness-based decision-making can shift into unconscious, conditioned patterns**.

These findings all point to a central idea: **Learning is not merely acquiring knowledge — it is the process by which awareness becomes encoded into patterns**, often freeing the mind for new challenges while delegating old ones to automatic routines.

The Decision-Making Process: From Awareness to Automation

Traditional models of decision-making emphasize rational analysis guided by beliefs, desires, and values[5]. But as Herbert Simon[6] argued, human decision-making is bounded by the mind's limitations —what he called *bounded rationality*. Because of these constraints, our decisions often rely on approximations, shortcuts, and intuition, rather than on exhaustive reasoning. Factors such as cognitive bias, emotion, upbringing, and even risk appetite influence the outcome of our choices, as elaborated in various models like expected utility theory[7][8].

From the perspective of the Trilogy Theory of Consciousness (TTC), awareness is the base for decision-making but on its own unfolds across four distinct stages—a framework that explains how intentional attention and Discretionary Selection of Intelligence for Awareness (DSIA) prepare the basis for free-willed, awareness-based decisions (Figure 1). In addition, as depicted in Figure 2, the process of decision-making is composed of three stages: preselection, selection, and post-selection appropriation[9]. These stages help us understand how the transition from free-willed, awareness-based decisions to habitual, automated ones occurs through the shift from Awareness-Based Choice Selection (ABCS) to Selection of

doi.org/10.32388/DLUQV1

Choice Based on Algorithm (SCBA), ultimately leading to Stimulus-Conditioned Behavioral Automation for repeated actions (Figure 3).

1. Preselection: The Stage of Mental Preparation

In TTC, the decision-making process begins with the **preselection stage**, where the mind prepares the groundwork for a decision by synthesizing a matrix of two core intelligences:

Informational Intelligence: Data drawn from sensory inputs, memories, acquired knowledge, beliefs, desires, and logical reasoning.

Emotional Intelligence: Signals from moods, physiological states (pain, hunger, fatigue), self-esteem, and affective memory.

These elements interact dynamically with foundational influences such as genetics, personal history, and neurobiological constraints. During this stage, the mind doesn't simply compute—it **reasons** and also **counter-reasons**, generating alternative narratives and outcomes. This dual processing is essential: without internal contrast and tension between options, true choice becomes meaningless.

This model extends and improves upon naturalistic decision models (e.g., [10]), which describe how goals are clarified and options evaluated. Yet those models often lack a mechanism for internal contradiction or counter-arguments—something TTC recognizes as fundamental to decision-making in natural intelligence.

TTC also reframes the conscious–unconscious debate. Where some theorists (e.g., [11]) attribute decision efficiency to unconscious mental capacity, TTC treats **the entire mind as unconscious** in its operational logic. Only when the **selection mechanism (ABCS)** is activated does awareness step in and tip the scale.

2. Selection: The Role of Awareness in Choosing

In the **selection stage**, the processed matrix of possible choices enters awareness—not all at once, but selectively, filtered through the mechanism TTC calls **Discretionary Selection of Information for Awareness**. This filtering explains why some aspects of a decision are vividly clear while others remain vague or inaccessible.

Here, **ABCS** comes into play. It is not simply rational calculation but an emergent moment when awareness influences the final decision. This means that the chosen option is not always the most logical, advantageous, or even goal-aligned—it is the one that resonates most within the momentary field of awareness.

Consider this example: You're looking for a job and ask a friend to go through classified ads and highlight potential roles. Your friend uses a logical, rule-based algorithm to flag every opportunity that matches your skills (this is SCBA—algorithmic selection). But as you scan the list, you instantly cross off a few options—say, a ballet dancer or nightclub bouncer—because your **awareness** of yourself renders those options meaningless. Your awareness didn't *compute* those eliminations; it simply *knew*. That shift—from logic to insight—is the defining mark of ABCS.

This selection mechanism, bounded by what DSIA allows into awareness, explains Simon's bounded rationality from a fresh angle. Rationality is not inherently limited—it is *filtered*, not fully disclosed to awareness at once.

3. From Selection to Automation: Enter SCBA

While ABCS reflects the free-willed, awareness-informed decision, repeated experience tends to shift decision-making toward **SCBA**. In SCBA, selection of choices is governed by **internalized algorithms**—patterns formed by past choices that now operate with minimal or no input from awareness. This is the hallmark of **learning** and **habit formation**.

Naturalistic models, such as those proposed by Klein[12] or Fox et al.[13], emphasize commitment and alignment with goals in decision selection. However, TTC goes further by distinguishing **automated commitment** (SCBA) from **conscious selection** (ABCS). While traditional theories might see repeated choice as deeper commitment, TTC shows it as a **migration from awareness selection of choice to automation**—from conscious deliberation to internalized pattern.

4. Post-Selection Appropriation: Encoding the Pattern

The third stage in TTC's decision-making model is **post-selection appropriation**—the process by which the selected choice, are appropriated for execution through being reinforced, modified, or discarded based on its outcome and subjective salience. This is where learning begins to take root.

Following the moment of decision (via ABCS), the mind monitors the **feedback** from that choice. This includes not only the anticipation or prediction of the external consequences (reward, punishment, success, failure), but also the **internal affective response**—satisfaction, regret, validation, doubt. These

doi.org/10.32388/DLUQV1

experiences are encoded into memory and gradually shape the matrix of informational and emotional intelligence for similar future decisions.

Critically, **repetition** of similar decisions in similar contexts leads to **neural and cognitive efficiency**. What was once a deliberative process becomes streamlined. Over time, the **awareness-driven act of choosing** is replaced by an **automated selection process**—a hallmark of **SCBA**. In this phase, awareness becomes increasingly bypassed. The input (stimulus) activates a familiar pattern of behavior (response) with minimal conscious involvement.

This transition is not a flaw—it is an **adaptive feature** of natural intelligence. It reflects the system's attempt to conserve energy and time by relegating known responses to internalized algorithms. It is how we form habits, develop skills, and create routines. It is also how conditioning works—turning **ABCS** into **SCBA** through feedback and repetition.

Bridging Into Conditioning and Learning

When viewed through this lens, classical conditioning (like Pavlov's experiment) and even complex human learning can be understood as a **gradual reorganization of the decision-making architecture**. What begins as an exploratory, awareness-based process (ABCS) becomes a refined, efficient algorithm (SCBA) once the environment proves that this pattern works.

One may consider the Pavlov's experiment was a compiled events of ABCS and an internal visceal reflex when the food result in salivation. In this scenario, the mind already had an inherent SCBA of secreting Saliva at the site of the food which was expanded to more complex form of SCBA. During the process the dog's first few salivations in response to the bell were likely preceded by confusion, attentiveness, and awareness. But after several pairings, the response became immediate, unthinking, automatic—**an SCBA embedded through repeated ABCS**. This same principle applies whether it's training a goldfish to gather at the surface or teaching a child to look both ways before crossing a street.

Thus, TTC reframes conditioning and learning not as mechanical operations, but as **evolutions of conscious engagement**, where **free-willed choice** is the raw material from which **automation** is forged.

The Purpose of Repetition: Efficiency Through Internalization

Repetition is not merely a mechanical loop—it is the engine of learning. In the context of TTC, each repeated choice made through **ABCS** contributes to a gradual migration toward **SCBA**. This shift reflects a

doi.org/10.32388/DLUQV1

profound evolutionary strategy: to improve efficiency and speed of an action **through a process of automation**.

The human mind is not designed to make every decision consciously. Doing so would be mentally exhausting and biologically inefficient. From tying shoelaces to driving familiar routes or speaking native languages, much of our daily functioning is automated. But this automation did not appear overnight—it was built upon countless moments of **conscious engagement**, each one strengthening the behavioral circuit.

In this perspective, **learning** is the process of **encoding awareness-based patterns into algorithmic structures**. This encoding allows behaviors to become faster, more reliable, and less demanding on the cognitive skills, reserving the awareness and cognition for more meaningful of novel tasks. Importantly, this is not a loss of intelligence—it is a **redistribution** of cognitive resources.

In TTC, awareness serves as a **limited but precious resource**—used to evaluate new, uncertain, or conflicting choices. Once a decision becomes validated by outcomes and reinforced over time, awareness can step back. What remains is an internalized algorithm: a behavioral routine that is fast, predictable, and efficient—**SCBA**.

Conditioning as Directed Rewiring

This process—when intentionally guided—is what we call **conditioning**. Whether in Pavlov's bell experiment, Skinner's operant conditioning, or modern behavior modification therapies, the aim is to **pair a stimulus with a desirable behavior** through repeated exposure and reinforcement.

Conditioning, then, is not simply external programming—it is a **rewiring of the decision architecture** from the inside out. The subject begins with ABCS: the initial exposure to a stimulus prompts awareness, evaluation, and selection. But as the contingencies become clearer, and as feedback loops tighten, the behavior becomes automatic.

In **classical conditioning**, the pairing of stimulus and outcome facilitates the **prediction** of reward or punishment.

In **operant conditioning**, the behavior itself is modified by its consequence—strengthened through reward or weakened through punishment.

In both forms, the endpoint is the same: behavior that was once tentative becomes **reliably triggered by the stimulus alone**, with no need for volitional reflection.

From the perspective of TTC, this is not a reduction of mind to reflex—it is **evidence of the mind's ability to optimize its use of awareness**. What begins as choice becomes protocol. What begins in flexibility becomes encoded as form.

The Double-Edged Sword of Automation

The transition from **ABCS** to **SCBA** represents a major cognitive achievement. It reflects the system's ability to optimize mental bandwidth, allowing awareness to disengage from familiar or repetitive decisions so it can be reserved for more complex, uncertain, or novel situations. But like all evolutionary solutions, automation comes with trade-offs.

Benefits of Automation: Efficiency and Stability

Cognitive Efficiency. Once a behavior is encoded as an SCBA, it requires **minimal cognitive effort**. This allows awareness to focus on higher-level tasks such as planning, innovation, or moral reasoning, while routine behaviors operate smoothly in the background.

Speed and Responsiveness. Automated behaviors are **faster** and often more **accurate** in familiar contexts than deliberated choices. They allow individuals to act swiftly in environments that reward quick, consistent responses (e.g., athletes, emergency responders, or skilled technicians).

Emotional Relief. Repetitive conscious engagement with emotionally charged decisions can be draining. Automating decisions in emotionally neutral or resolved domains provides **psychological relief**, preventing decision fatigue.

Behavioral Consistency. SCBA provides **stability**—a structured behavioral identity shaped by one's past experiences and refined through reinforcement. This helps establish reliable habits, routines, and even social predictability.

Costs of Automation: Rigidity and Blind Spots

Loss of Flexibility. While SCBA is efficient, it is **not adaptive** in changing or novel environments. Automated behaviors can persist even when they're no longer beneficial—what was once a useful habit can become an **obstacle to growth**.

Bypassing Awareness. When behavior becomes conditioned, awareness may not even **register the act**. This can result in **mindless repetition**, where individuals act out patterns without fully understanding why, or without reflecting on whether the choice aligns with current goals.

Entrenched Biases and Maladaptive Habits. Conditioning can encode not only helpful routines but also **dysfunctional or harmful behaviors** (e.g., avoidance responses, compulsions, addictive patterns). Once internalized, these patterns become difficult to change, as they bypass ABCS.

Reduced Sense of Agency. Perhaps most significantly, excessive reliance on SCBA can erode the **experience of free will**. When many decisions are made below the threshold of awareness, individuals may feel **disconnected from their own choices**, caught in a loop of behavior that no longer feels authored by the self.

Rethinking Libet's Experiment: SCBA Misread as ABCS

One of the most influential challenges to the notion of free will emerged from the work of Benjamin Libet[1], who demonstrated that a brain signal—termed the *readiness potential*—preceded the subject's reported awareness of intending to move. His conclusion was provocative: if the brain initiates an action before conscious intention arises, then free will must be an illusion.

But from the perspective of TTC, Libet's interpretation may rest on a category error.

Libet assumed that the act of intention he was measuring corresponded to a conscious, volitional decision—what TTC defines as **ABCS**. However, closer scrutiny suggests that the experimental setup itself may have favored **automation**, not volition. Participants were instructed to perform a simple, repetitive movement (e.g., flicking a wrist) at arbitrary moments. With each repetition, the decision to act likely shifted from ABCS toward **SCBA**—the automated execution of behavior that no longer requires active awareness for initiation.

In such a setup, the **awareness of the selected choice** naturally lags behind the neural preparatory signal—not because awareness lacks causal power, but because the action has already become conditioned. What Libet measured, then, was not the initiation of a free choice, but rather the **byproduct of a learned behavior loop**.

TTC clarifies that **intention arises from ABCS**, but that **awareness itself is not synonymous with intention**. Rather, awareness initiates the process through intention (i.e., ABCS), and then remains engaged through a series of **feedback loops** that sustain awareness of the selected choice, its appropriation, and eventual execution. This ongoing engagement of awareness (meta-awareness) is distinct from the **initial volitional act**—which Libet mistakenly identified as the delayed conscious signal.

doi.org/10.32388/DLUQV1

In this light, the temporal gap Libet observed—between readiness potential and awareness of the selected choice—simply indicates that **the intention (ABCS) was no longer involved** in the conditioned behavior, rather than disproving free will. It reflects a **misidentification of SCBA as volitional choice**.

Therefore, the conclusion that free will is an illusion may stem from a misunderstanding of **which stage of decision-making** was being observed. Libet captured the outcome of a behavioral circuit that had already been automated. Had his study been designed to track **novel, unconditioned decisions**—true ABCS engagement—the sequence of neural and conscious events might have looked very different.

Balancing Awareness and Automation

TTC does not present SCBA as a flaw—it is a **necessary evolutionary feature**. But the theory also highlights the importance of **meta-awareness**: the ability to occasionally step back and re-engage awareness even in domains dominated by SCBA. This is the **key to behavioral change**, personal growth, and the reconfiguration of one's internal algorithms.

Just as awareness seeds the formation of habits, it must also retain the capacity to **reclaim them** when change is needed. In this way, TTC offers not just an explanation for how we automate our behavior—but a map for how to **rewire it**.

Future Directions: Investigating the Transition from Awareness to Automation

The TTC provides a novel lens to understand how choices evolve—from deliberate, awareness-based acts (ABCS) to automatic, algorithm-driven routines (SCBA). While philosophical and cognitive frameworks support this model, **experimental validation is the next logical step**. To empirically observe the shift from ABCS to SCBA, we must design studies that capture this **invisible transition zone**—the point at which awareness disengages, and behavior becomes internally coded.

One proposed experiment draws from the domain of **motor learning and procedural memory could involve the tasks of p[laying music. an experiment** to identify the transition point between conscious awareness and automated motor behavior during repetitive musical performance. For example, **how fast** a short music sheet consisting of **repeated note sequences** with structured variation, can be considered an automated process in regard to the speed of playing on a **digital keyboard** to record finger position, keystroke timing, and velocity to show signs of automation such as **reaction time per note**, **accuracy**, and

inter-key timing as proxies for effort. or using imaging technique showing neuroimaging (e.g., fMRI or EEG) to determine **neural correlates of the ABCS→SCBA transition**, such as the diminishing activation of prefrontal areas and increased basal ganglia involvement over time.

Conclusion

The reinterpretation of Libet's experiment through the lens of the Trilogy Theory of Consciousness reframes not only our understanding of volition but the broader architecture of learning and conditioning. Rather than proving free will to be an illusion, Libet's findings may have inadvertently measured a behavior already delegated to automation—a case of SCBA mistaken for a conscious act of will. TTC proposes that awareness plays a critical initiating and regulatory role in decision-making, but one that can be gradually bypassed as behaviors are internalized.

By distinguishing between awareness-driven intention (ABCS) and algorithmic automation (SCBA), TTC clarifies the layered nature of cognition—where conscious agency gives rise to habits, and meta-awareness can reclaim or reprogram them. This framework not only resolves the apparent conflict between neuroscience and subjective experience, but also offers a roadmap for intentional behavioral transformation. If awareness can seed automation, it can also revisit it. The mind, though wired by experience, remains rewritable by intention.

References

1. ^a, b^ *Libet B, Gleason CA, Wright EW, Pearl DK (1983). "Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act." Brain. 106(3):623–642. doi:10.1093/brain/106.3.623.*
2. ^ *Johansen JP, Tarpley JW, LeDoux JE, Blair HT (2010). "Neural substrates for expectation-modulated fear learning in the amygdala and periaqueductal gray." Nature Neuroscience. 13(8):979–986. doi:10.1038/nn.2594.*
3. ^ *Takahashi YK, Batchelor HM, Liu B, Khanna A, Morales M, Schoenbaum G (2017). "Dopamine Neurons Respond to Errors in the Prediction of Sensory Features of Expected Rewards." Neuron. 95(6):1395–1405.e3. doi:10.1016/j.neuron.2017.08.025.*
4. ^ *Doll BB, Duncan KD, Simon DA, Shohamy D, Daw ND (2015). "Model-based choices involve prospective neural activity." Nature neuroscience. 18(5):767–772. doi:10.1038/nn.3981.*

5. ^Slovic P, Fischhoff B, Lichtenstein S (1977). "Behavioral decision theory." Annual Review of Psychology. 28: 1–39. doi:10.1146/annurev.ps.28.020177.000245.
6. ^Simon HA (1956). "Rational choice and the structure of the environment." Psychological Review. 63(2):129–138. doi:10.1037/h0042769.
7. ^Frisch D, Baron J (1988). "Ambiguity and rationality." Journal of Behavioral Decision Making. 1(3):149–157. doi:10.1002/bdm.3960010303.
8. ^Steele K, Stefansson HO (2020). "Decision theory." In Zalta EN (Ed.), The Stanford Encyclopedia of Philosophy (Fall 2020 Edition). https://plato.stanford.edu/archives/fall2020/entries/decision-theory/.
9. ^Farhadi A (2023). "Trilogy: A New Paradigm of Consciousness." Neuropsychiatry (London). 13(4):1–16.
10. ^Drummond H (1991). Effective Decision Making: A Practical Guide for Management. Springer.
11. ^Dijksterhuis A (2004). "Think different: The merits of unconscious thought in preference development and decision making." Journal of Personality and Social Psychology. 87(5):586–598. doi:10.1037/0022-3514.87.5.586.
12. ^Klein G (2015). "A naturalistic decision making perspective on studying intuitive decision making." Journal of Applied Research in Memory and Cognition. 4(3):164–168. doi:10.1016/j.jarmac.2015.07.001.
13. ^Fox CR, Rogers BA, Tversky A (2013). "Commitment to a course of action." In Lichtenstein S, Slovic P (Eds.), The construction of preference. Cambridge: Cambridge University Press. pp. 181–200.

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.