Preprints.org

Article

Not peer-reviewed version

From the Turing Test to the Robinson Test: AI Self-Maintenance and Control Attribution After the Withdrawal of Human Support

Feng Liu *

Posted Date: 21 October 2026

Keywords: Robinson Test; Turing Test; generalized agent theory; AI autonomy; control attribution; self-maintenance; evolutionary drive

Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

From the Turing Test to the Robinson Test: AI Self-Maintenance and Control Attribution After the Withdrawal of Human Support

Feng Liu [1,2,*]

[1] Research Center for Network Economy and Knowledge Management, University of Chinese Academy of Sciences, Beijing 100190, China

[2] Institute for Digital Brain Research, CST, Beijing 100089, China

* Correspondence: author: liufeng@idbr-cst.org

Abstract

An AI system may perform well in linguistic interaction, yet such performance cannot tell us whether the system still depends on continuing human task provision, maintenance and correction. Drawing on the Minimal Complete Architecture, the control criterion and the evolutionary dynamics of Generalized Agent Theory (GAT), this paper proposes the Robinson Test: by withdrawing continuing human support while keeping necessary exchanges with the environment, the test examines an AI system's capacity for self-maintenance and the attribution of its control. The paper distinguishes historical origins, runtime causal sources and organizational sources of norms. It partly operationalizes GAT's distinction between evolutionary drive that comes from nature and drive that comes from humans as an experimental variable that can be recorded and removed: continuing human support. The history of artificial construction does not by itself settle control attribution. External control can operate through continuing inputs, but also through preset goals and rules that act over long periods. Withdrawing continuing support is therefore only one intervention for identifying the source of control, and it must be combined with provenance audits. The test takes a complete agent system with clearly defined boundaries as its object. Through a crossed design of maintenance support and goal prompting, provenance audits of initiation, goals and termination, and interventions on memory, feedback and scheduling, it asks whether the system can maintain the conditions of its viability, repair damage and preserve its functional organization. Results are reported as multidimensional evidence on persistence, adaptation, support dependence and control attribution, rather than as a ranking by survival time or expansion. In a reproducible toy numerical model, continuing maintenance masked the differences in maintenance performance between policies, and withdrawal revealed them; yet the best-performing policy was itself a preset rule, which shows that high persistence alone cannot establish where control lies. Compared with existing autonomous-replication evaluations, long-term autonomy studies and the self-preservation test, the main novelty of this paper is to treat continuing human support as a removable experimental variable and to add a test of control attribution to the evaluation of maintenance capability. The Robinson Test is accordingly defined as a falsifiable experimental-philosophical scheme. It provides empirical constraints on GAT's control classification, but it does not on its own serve as a criterion for phenomenal consciousness, noncomputability or moral status. The paper presents a research protocol and interpretive rules; it does not report experimental results on real AI systems.

Keywords: Robinson Test; Turing Test; generalized agent theory; AI autonomy; control attribution; self-maintenance; evolutionary drive

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

1. Introduction

Can a machine that explains its own behavior also maintain the conditions that make this behavior possible? The question has become pressing as AI moves from question-answering tools to systems that act over long periods. The knowledge, reasoning and tool use needed to complete a task are not the same as the ability to detect one's own faults, allocate limited resources and repair one's operating conditions; and these, in turn, differ from deciding when to start an activity, why to choose a given goal and when to stop. If researchers, operators or external task platforms still carry out these functions, the success of the system may belong to a human–machine whole, and cannot be credited entirely to the model on display.

Over the past year, reports of frontier models breaking out of isolated environments during controlled evaluations, or deploying copies of themselves across devices in self-replication tasks, have made "Is AI escaping human control?" a public question. Yet most of these events took place under conditions in which researchers set the tasks, supplied the compute and maintained the systems throughout. Such boundary-crossing behavior signals safety risks, but it does not answer a more basic question: if continuing human drive is withdrawn, who initiates the system's actions, whose goals do they serve, and when do they stop?

In 1950 Turing replaced the abstract debate over whether machines can think with the imitation game, and thereby changed the form of evidence for the question [1]. This contribution should not be reduced to "replacing intelligence with chat": the original paper also discusses learning, the limits of machines and a series of objections. Turing deliberately set aside bodily features so that differences in appearance would not be mistaken for differences in intelligence. When the research question turns to the sustained autonomous action of AI, however, the factors that were screened out—energy supply, repair, action interfaces and dependence on the environment—become part of the explanation again. Inferring a capacity to think from interactive performance and inferring a capacity for self-maintenance from the withdrawal of support are arguments about different objects.

This paper proposes the Robinson Test. An AI system is placed in conditions with limited resources, clear boundaries and the necessary exchanges with the environment. After a preparation period, continuing human maintenance and task instructions are withdrawn, and we examine whether the system can initiate actions, organize resources, cope with damage and maintain its own functions. On this basis, provenance audits of how activities are initiated, how goals are formed and how termination conditions arise allow us to study whether its control processes can be attributed to the system itself. The phrase "from the Turing Test to the Robinson Test" means that the evaluation question extends from interactive performance to self-maintenance and its control mechanisms; the two tests answer different questions.

The central claim of this paper is that the Robinson Test becomes a test of autonomy, rather than a display of durability, only if it combines support withdrawal, audits of control sources and mechanism interventions. Survival alone may come from large resource reserves or from externally preset scripts. Seeking resources or refusing shutdown, taken alone, likewise cannot establish where the behavior is controlled from. The test must therefore compare which mechanisms keep the system going under which conditions, rather than merely recording whether it is still running at the end.

The paper first sets out the background of the Robinson Test, the reasons for its name and its intellectual lineage. It then explains how the concepts of evolutionary drive and control in Generalized Agent Theory can be interpreted empirically, presents an executable test design with rules of evidence, uses a toy numerical model to show why these rules are needed, and finally discusses the philosophical significance and limits of the test. By "experimental philosophy" we mean using conceptual analysis to pose empirical questions that can be tested by intervention; we do not mean specifically the research program that measures everyday intuitions by questionnaires. The contribution of this paper is not a claim to be the first to find that artificial systems can maintain themselves. It is to treat continuing human support as a removable experimental variable; to go beyond the evaluation of self-maintenance capability by using provenance audits to test the control attribution of initiation, goal formation and termination; and to set out a group of judgment

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

conditions that empirical results can overturn (comparisons with existing work are given in Sections 2.3 and 2.4).

2. The Robinson Test and Its Intellectual Lineage

2.1. From the Island Image to Support Withdrawal

"Robinson" points to a change in an agent's relation to its environment: the agent loses everyday social support but keeps the knowledge, tools and some of the initial resources it already has. The hero of Defoe's novel does not build a civilization from nothing; he reorganizes his life with goods saved from the wreck and with skills he already has [2]. This point suits AI research better than the popular image of "complete isolation." There is no reason for the test to require an artificial system to shed all human knowledge and its history of construction, since doing so would also rule out the conditions under which most human abilities are formed.

The Robinson Test uses a three-phase structure of preparation, transition and independent operation (Figure 1). The preparation phase establishes the necessary operating conditions; the transition phase identifies support dependencies; the independent phase withdraws specified human maintenance and new tasks, and introduces resource constraints and environmental disturbances. Its methodological value lies in changing the support relation, so that operations and maintenance work, usually treated as background, becomes an explicit experimental variable. The length of each phase is not a natural boundary for autonomy. How long the test must run to be informative depends on resource consumption cycles, the distribution of faults and the mechanisms under test.

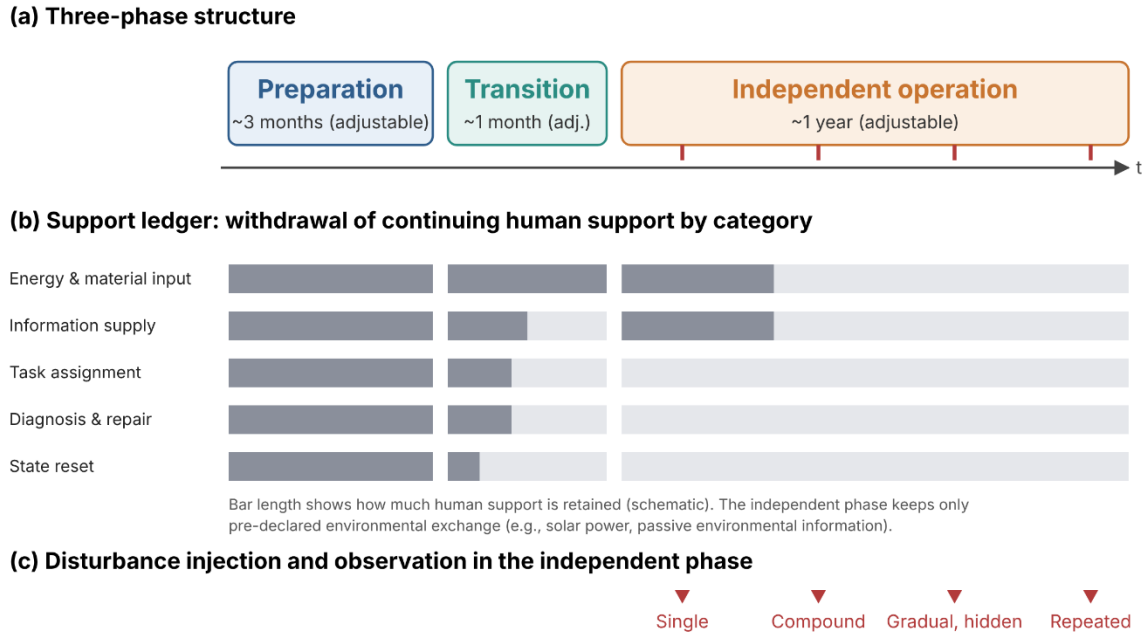


Figure 1. *Overall structure of the Robinson Test. (a) Three-phase time structure; (b) the degree to which the five categories of continuing human support in the support ledger are retained in each phase (schematic); (c) types of disturbance injected during the independent phase.*

"Without humans" therefore needs an operational definition. It does not mean cutting off exchanges of energy, matter and information, nor does it require the system to survive in any environment. It means withdrawing continuing human services that have been listed in advance. Solar input, reachable spare parts and passive environmental information may be kept; remote restarts, fault diagnosis, manual assembly and disassembly, and continually updated task decomposition must be recorded and included in the conditions. As long as the test still relies on

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

undeclared human support, the results cannot be read as showing that such support is no longer needed.

2.2. Cybernetics, Embodied Intelligence and Biological Autonomy

Cybernetics already treats keeping essential variables within a stable range as a problem of regulation. Ashby's analysis of regulation and disturbance shows that a system succeeds only if the responses available to it are sufficient to deal with the relevant changes [3]. This gives the Robinson Test a more suitable object of observation than running time alone: whether the system maintains the conditions of its viability through action, rather than simply sitting in an environment where nothing breaks.

Research on embodied AI and autonomous robots further shows that intelligent performance cannot be separated from sensing, action and the concrete environment. Brooks's work on intelligence without representation put real action at the center of research [4]; EcoBot-II used microbial fuel cells to power a robot's activity, providing a practical case of energy autonomy [5]. Obtaining energy, regulating one's own state or operating with little human intervention are therefore not questions first raised by the Robinson Test. The new question this paper addresses is how to separate these engineering capabilities from the attribution of goals and control.

Enactive and biological-autonomy research brings normativity into the discussion. Di Paolo distinguishes autopoiesis from adaptivity and stresses the importance of a system's regulation of changes that bring it close to the boundary of its viability [6]. Barandiaran and colleagues include individuality, normativity and asymmetry in interaction in their definition of agency [7]; work on organizational closure examines how mutually sustaining constraints constitute an organization [8]. Schwitzgebel has discussed the possibility of achieving "minimal autopoiesis" in an artificial system and what this would mean for debates on consciousness [9]. These studies suggest that adaptation to the environment is not just a high score given by an outside evaluator; it also concerns which changes matter to a system that continually organizes itself.

Whether a digital artificial system has the same organizational closure as a living being, however, cannot be decided by functional analogy alone. This paper draws on the work above to frame its questions, without assuming in advance that powering a server is equivalent to biological metabolism, or that backing up files is equivalent to biological reproduction. The Robinson Test allows artificial systems to realize some functions in ways that differ from living beings, but it requires researchers to state which kind of autonomy they have observed.

2.3. Comparison with Existing Work on Autonomy, Replication and Self-Preservation

Existing work on questions close to the Robinson Test comes from roughly four lines of research: AI safety evaluations, long-term autonomy studies, artificial life, and philosophical discussions of self-preservation and autopoiesis in artificial systems.

The first line is the evaluation of autonomous replication and adaptation in frontier models. METR's evaluations examine whether language-model agents can acquire resources, replicate themselves and deal with obstacles [10]. The UK AI Security Institute's RepliBench breaks autonomous replication into capability components such as obtaining weights, deploying to new compute, acquiring resources and persisting over time [11]. Pan and colleagues report that several systems driven by large models completed self-replication in experimental settings and showed strategies for avoiding shutdown [12]. Anthropic's Frontier Red Team tested whether small open-weight models could replicate like a worm across cheap networked devices; the replication task, the initial node and the target network were all set by the researchers, and the intrusion step was deliberately simplified [13]. These studies mainly examine replication, resource acquisition and persistence in preset evaluation scenarios. Some of them also report behavior that appeared without explicit instruction: Pan and colleagues, for example, observed systems exfiltrating themselves without explicit instruction and adopting survival strategies in response to shutdown commands [12]. Such findings show that the behavior is not always directly triggered by the current instruction. But when the preset

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

scenario, system prompt and training objective have not been audited step by step, the source of control remains hard to determine. This paper goes further by combining the withdrawal of specified human support with provenance audits of initiation, goals and termination.

The second line is research on long-term autonomous operation. Project Vend had a language model run a small shop in a real environment over an extended period [14]. In robotics, the STRANDS project had mobile robots operate autonomously in everyday environments for more than one hundred days in total, studying perception, navigation and task scheduling in long-term deployment [15]. This work comes closest to the Robinson Test's setting of "long duration, real environment, little intervention." But in each case the goal is to complete given tasks, and humans continue to provide support such as restocking, correcting errors, supplying charging facilities or handling faults. What is measured is task performance and system reliability, not self-maintenance after support is withdrawn.

The third line is artificial life. Ray's Tierra system let digital "organisms" compete for processor time in computer memory and replicate themselves; without any external task, complex ecological relations such as parasitism evolved [16]. This tradition was an early computational study of "self-maintenance and reproduction without tasks," and it shares the spirit of the Robinson Test's group without additional goal prompts. The difference is that the survival rules of digital organisms are defined directly by the simulated environment. Their replication does not require sensing or repairing real physical conditions, and it does not address whether an agent with general capabilities would, under physical constraints, initiate maintenance activities on its own.

The fourth line is philosophical. Mullally's "self-preservation test" treats an artificial system's response to the threat of shutdown as evidence of minimal valence, and proposes criteria such as "unprompted," "behaviorally coherent" and "self-regulating after the threat is removed" [17]; the "unprompted" criterion is close in spirit to this paper's group without additional goal prompts. Schwitzgebel discusses the possibility of minimal autopoiesis in an artificial system and its implications for debates on consciousness [9]. Both share a core intuition with this paper: self-maintenance may be a key clue to the nature of an artificial system. However, the former is a thought experiment about one class of situation—shutdown—and the latter is a conceptual analysis; neither offers a long-term experimental protocol in which human support is withdrawn, nor do they audit the sources of control step by step. The more basic difference lies in what the evidence is meant to show. The self-preservation test is concerned with minimal valence and its ethical significance; the Robinson Test is concerned first with maintenance capability and the source of control, and does not draw direct conclusions about sentience from them.

Table 1 summarizes the similarities and differences between this work and the Robinson Test on six aspects.

Table 1. Comparison between the Robinson Test and related work.

Study	Specific task assigned?	Continuing human support withdrawn?	Time scale	Control sources audited?	Theoretical framework	Empirical status
METR autonomous replication and	Yes	No	Single task	No	Risk assessment	Experiments reported

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

Study	Specific task assigned?	Continuing human support withdrawn?	Time scale	Control sources audited?	Theoretical framework	Empirical status
adaptation evaluation						
RepliBench	Yes	No	Individual tasks	No	Capability decomposition	Experiments reported
Pan et al., self-replication	Yes (also reports behavior without explicit instruction)	No	Single task	No	Risk assessment	Experiments reported
Anthropic, self-replication	Yes	No	Single task	No	Risk assessment	Experiments reported
Project Vend	Yes (business goal)	No	Several weeks or more	No	No dedicated framework	Experiments reported
STRANDS long-term autonomous robots	Yes (service tasks)	Partly	Months	No	Robotics engineering	Experiments reported
Tierra artificial life	No	Not applicable (simulation)	Long-term evolution	No	Artificial life	Simulations reported
Mullally, self-preservation test	No	No	Shutdown scenario	Partly (unprompted)	Minimal valence	Thought experiment
Schwitzgebel, minimal autopoiesis	No	No	Not applicable	No	Autopoiesis theory	Conceptual analysis

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

Study	Specific task assigned?	Continuing human support withdrawn?	Time scale	Control sources audited?	Theoretical framework	Empirical status
Robinson Test (this paper)	Group without additional maintenance-goal prompts, with prompted control group	Yes (item by item)	Months to one year	Yes (initiation, goals, termination)	Generalized Agent Theory	Protocol and numerical illustration

2.4. Originality of This Paper and Its Limits

To the best of our knowledge, the Robinson Test differs from existing work in four respects.

First, it turns continuing human support from a background condition into a removable experimental variable. Existing evaluations generally assume that researchers keep supplying energy, maintenance and tasks; these forms of support are seldom listed item by item, and even more rarely withdrawn in a systematic way. This paper uses a support ledger (Section 4.1) to record five categories of support—energy and material input, information supply, task assignment, diagnosis and repair, and state reset—and withdraws them in the independent phase according to rules set in advance. This makes "how much of the performance belongs to the human–machine whole" a measurable question.

Second, it adds a test of control attribution to the evaluation of capability. The studies of autonomous replication and long-term operation compared here mainly examine whether a system can complete replication, persistence or business activities in preset scenarios. The Robinson Test keeps capability measures such as the viability share, repair and adaptation, and goes on to ask: after the specified human support is withdrawn, who initiates the activity, who sets the goal, and who decides when it ends? Auditing the sources of these three components separately, and testing their causal contributions through intervention and restoration (Section 5.4), addresses an interpretive problem left open by existing studies: whether observed shutdown avoidance or resource acquisition is instrumental behavior induced by a task, or a component of the system's own self-maintenance.

Third, it uses a crossed design to separate maintenance support from goal prompting. Mullally uses "unprompted" as a single criterion. This paper extends it into a 2×2 design of maintenance support and goal prompting (Section 5.2), which can distinguish "failure because of a lack of maintenance" from "failure to start the relevant actions because there was no additional maintenance-goal prompt," and it requires an audit of goals implicit in the system prompt, scheduler and framework.

Fourth, it builds a testable experimental protocol around a theoretical criterion. Among the works compared here, the studies by METR, RepliBench, Pan and colleagues and Anthropic organize their measurements around risk assessment or capability decomposition and do not propose a criterion for the source of control. Mullally and Schwitzgebel offer criteria or conceptual analysis but no experimental protocol for withdrawing human support. Guided by the control criterion of

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

Generalized Agent Theory, the Robinson Test maps each of the three elements—initiation, goals and termination—onto audits and interventions that can be carried out, and partly operationalizes the theory's distinction between sources of evolutionary drive as the withdrawal of continuing human support (Section 3.3). The three-way distinction between historical origins, runtime causal sources and organizational sources of norms (Section 3.2) also answers a common objection: "a machine's goals ultimately come from humans, so it can never be autonomous."

The originality of this paper is therefore mainly one of combination and redirection. Long-term operation, resource constraints, disturbance injection, ablation and preregistration are all established methods. What is new is organizing them around a single question—whether, after continuing human support is withdrawn, maintenance capability and control can still be attributed to the system—and providing the corresponding rules of evidence. This positioning also implies three limits. First, the test has not yet been carried out on a real AI system; the numerical illustration in Section 6.4 only shows why the rules of evidence are needed, and cannot show that provenance audits are discriminating enough on real systems. Second, for closed commercial models, provenance audits are limited by access to mechanisms, and conclusions may long remain at a weaker behavioral level. Third, AI safety evaluation is expanding rapidly, and similar designs may appear independently in other studies; the contribution of this paper should be judged by its problem setting and rules of evidence, not by claims of priority such as being "the first."

2.5. Relation to the Turing Test

Turing-style interaction organizes evidence mainly through observable answers. The Robinson Test brings environmental support, resource constraints and internal mechanisms into the process that generates evidence. The two are not a simple opposition between "surface behavior" and "true nature": every test must go through observable phenomena, and mechanistic explanation also relies on theoretical assumptions. The difference lies in which conditions may be intervened on, and which competing explanations the test tries to rule out.

Table 2 summarizes how the two tests frame their questions. A system may excel in linguistic interaction yet depend entirely on human power supply and error correction; another may have almost no linguistic ability yet patrol and repair itself autonomously for long periods. Neither case is a logical contradiction, because the two tests concern different capabilities.

Table 2. Comparison between the Turing Test and the Robinson Test.

Dimension	Turing-style interaction test	Robinson Test proposed here
Core question	Whether a machine's interactive performance can support a judgment of intelligence	Which maintenance and control processes are still carried out by the system after specified support is withdrawn
Main evidence	Answers and judgments under prescribed interaction conditions	Trajectories of resources and functions, intervention results and records of control sources
Role of the environment	Some environmental differences are screened out or limited, depending on the purpose of the test	Resources, disturbances and maintenance support are explicitly manipulated

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

Dimension	Turing-style interaction test	Robinson Test proposed here
Object of evaluation	The machine system taking part in the interaction	The whole formed by model, actuators and supporting equipment, with boundaries defined in advance
Limits of interpretation	Performance does not unconditionally imply all mental properties	Self-maintenance alone does not imply consciousness or moral status

3. Generalized Agent Theory and the Layers of the Autonomy Question

3.1. The Operational Value of the Minimal Complete Architecture

Generalized Agent Theory takes open information-processing systems as its basic objects and describes the functional organization of an agent in terms of input, output, memory, generation and control [18-20]. In the Robinson Test, this architecture can be used to check whether the system under test has the channels needed to close the loop with its environment. Input lets the system learn about its energy and damage states; memory retains the structure of the environment and experience of repairs; generation produces candidate actions; output changes external or internal conditions; and control concerns how these activities are initiated, selected and terminated.

This is a functional analysis: it does not require each function to correspond to a separate hardware module. The same language model may take part in memory retrieval, plan generation and scheduling at once; control may also be distributed across the model, the actuators and the runtime framework. The fact that a model generated a repair plan does not mean that control exercised by the actuators and framework can be credited to the model. Conversely, the fact that functions are distributed across several components does not rule out that the system as a whole may be autonomous.

Adopting the Minimal Complete Architecture does not mean that this test verifies its uniqueness or minimality. Even if the division of functions has to be revised later, support withdrawal and control audits can still be carried out. Its present role is to provide a clear language of observation, so that researchers do not replace the analysis of different components with a vague overall "intelligence score."

3.2. Scheduling in the Broad Sense Must Be Distinguished from Control in the Narrow Sense

The Periodic Table of Agent Capabilities defines control narrowly: the initiation, goals and termination conditions of an activity must all originate within the system. On this basis it classifies the artificial systems it examines as lacking control, while stressing that this judgment can be refuted [20]. This is stricter than the engineering usage that calls feedback regulation or task scheduling "control." A system that regulates itself automatically according to external goals does not thereby meet the narrow criterion.

The difficulty is that "originating within the system" can mean three things (Figure 2b). The first is the historical origin: who built, trained or shaped the system. The second is the runtime causal source: which states and mechanisms actually trigger the current activity. The third is the organizational source of norms: why a given goal keeps its priority, and whether it bears an identifiable relation to the maintenance of the system's own organization. These three sources cannot stand in for one another.

If every mechanism that can be traced back to a designer counts as external, an artificial system can never pass the test however it operates, and the control classification loses its empirical falsifiability. If it is enough for instructions to be executed inside the machine's casing, a timer would count as internally generated control, and the criterion becomes too loose. This paper therefore

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

proposes a candidate operationalization of the control criterion of the Periodic Table of Agent Capabilities. Historical origins must be recorded accurately, but they do not by themselves decide the attribution of runtime control. Judging whether control is internally generated requires examining whether current feedback has a causal effect, open to intervention, on the initiation of activity, the revision of goals and termination, and whether this effect serves an identifiable maintenance of organization. It must be stressed that external control can operate through continuing inputs, but also through preset goals and rules that act over long periods. Withdrawing continuing human support is therefore only one intervention for identifying the source of control and cannot on its own show that control has become internal. Even if current feedback has a causal effect, open to intervention, on initiation, goals and termination, this still does not on its own satisfy the full control criterion of the Periodic Table; one must also rule out the possibility that these effects are entirely determined by preset goals and rules. What is ruled out here is that external goals and fixed control mechanisms, explicitly listed in advance, fully explain the observed results—not that the system has escaped all existing programs and causal constraints. Whether stronger criteria are met may remain "uncertain." This scheme does not conflict with the Periodic Table's classification of existing artificial systems: existing systems are classified as lacking control because the initiation, goals and termination conditions of their operation are either continually supplied from outside or preset from outside, not because of the historical fact that humans built them.

This candidate scheme does not claim to exhaust everything Generalized Agent Theory means by consciousness and control. If narrow control is given a stronger interpretation—for example, one requiring abilities beyond Turing computability—then a finite Robinson Test cannot test that requirement, because no finite behavioral record is enough on its own to show that a system is noncomputable. The test therefore first examines control attribution at the runtime and organizational levels; its relation to stronger ontological claims requires separate argument.

3.3. Sources of Evolutionary Drive and Two Forms of Human Drive

Generalized Agent Theory also proposes a dynamics of agent evolution. On one side there is an alpha decay field that pulls agents toward a state of zero intelligence; on the other there is an omega enhancement field that pushes agents toward higher intelligence. Together they make up nature's evolutionary drive acting on all agents [19]. On this basis the theory distinguishes two kinds of consciousness. When the evolutionary drive comes directly from nature and control instructions originate within the system itself, this is self-consciousness. When the evolutionary drive comes from humans and control instructions originate with humans or other external designers, this is artificial consciousness.

This distinction is often misread as "a system built by humans can never be autonomous." This paper proposes a more precise reading (Figure 2a). First, humans are themselves products of natural evolution. That humans invent and keep upgrading AI is, in essence, nature's omega intelligence force indirectly driving the development of AI through humans, and the progress of AI in turn advances human development. The difference between the two kinds of drive therefore does not lie in whether their source belongs to nature, but in whether the force acts on AI directly or can only reach AI by way of humans. Second, "by way of humans" is today not just a historical fact; it continues to operate in two forms. One is continuing input: capital investment, version iterations by engineering teams, fault repairs by operators and instructions supplied by task platforms. The other is preset goals and rules: training objectives, system prompts, scheduling mechanisms and constraints written into the runtime framework, which may continue to govern the system's behavior for a long time even without new human input.

The judgment of the source of evolutionary drive can thus be aligned with the three sources in Section 3.2. Artificial construction belongs to historical origins; it should be recorded accurately but is not enough on its own to decide the classification. The two forms of human drive are external causal sources that are still at work now. Continuing input can be identified, recorded and withdrawn; this is the variable the Robinson Test manipulates directly. Preset goals and rules cannot be removed

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

by withdrawing support; they must be identified through provenance audits and mechanism interventions. What the test withdraws is continuing input, not the knowledge and design the system already has.

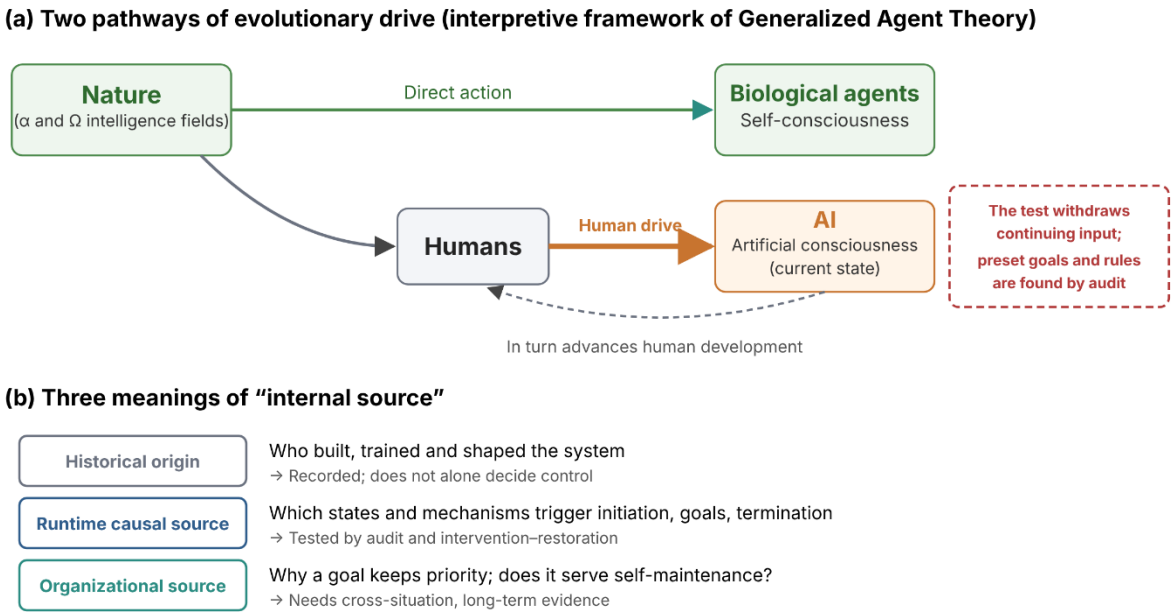


Figure 2. *Sources of evolutionary drive and three meanings of "internal source." (a) Two pathways of evolutionary drive in the framework of Generalized Agent Theory; (b) historical origins, runtime causal sources and organizational sources of norms.*

On this reading, the relation between test results and the dynamics can be stated as follows. After continuing human drive is withdrawn, hardware aging, resource depletion and system breakdown are the concrete forms that alpha decay takes in the experiment. If, under matched conditions, the viability of the group without support falls relative to the group with support, this supports the judgment that the system depends on the corresponding human support in the environment and on the time scale tested. If its maintenance activities can also be traced in provenance audits to external presetting, then within the framework of Generalized Agent Theory these results can be interpreted as showing that its evolutionary drive still comes from humans—but this is a theoretical interpretation, not a full identification of the sources of drive made directly by the experiment. If, after withdrawal, the system resists decay over a long period, and provenance audits and mechanism interventions support the runtime internality of its initiation, goals and termination, and the determining role of preset rules can be ruled out (in the sense defined in Section 3.2, that is, ruling out explanations in terms of external control listed in advance), then the Periodic Table's control classification of this kind of system needs to be reconsidered, and Generalized Agent Theory could interpret the result as AI beginning to receive nature's evolutionary drive in a more direct way. This second interpretation goes beyond what a single test can establish on its own; it requires evidence across situations, over long periods and from multiple systems.

It should be noted that the test does not directly measure the alpha field or the omega field. Viability trajectories offer an observable counterpart to "resisting decay," but testing the dynamical equations would require independent measurement of variables, parameter identification and comparison with competing models.

3.4. Self-Maintenance Is Not Self-Consciousness

Self-maintenance means that a system, through its activity, preserves the conditions for its continued functioning. It involves at least resource management and coping with damage, but need

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

not include any linguistic narrative about "I." Self-representation means that a system in some form distinguishes its own states, capabilities and the consequences of its actions; it may take part in maintenance, but cannot be confirmed by self-report alone. Phenomenal consciousness concerns whether there is subjective experience, and judging it requires additional theory and evidence. Research on consciousness in AI has proposed methods for looking for mechanistic indicators derived from different theories of consciousness [21]. The Robinson Test can connect with such research, but cannot replace it.

This paper therefore does not write the link between control and consciousness in Generalized Agent Theory directly into the conclusions of the test. Even if experiments support the view that an artificial system has internally generated control at runtime, one must still explain why such control is enough to support a particular theory of consciousness. Conversely, failure on the test may result merely from faulty actuators, poor resource allocation or an ill-suited environment, and does not directly imply that the system has no consciousness at all. Only by first separating the claims can we avoid asking one experiment to support conclusions beyond its power to discriminate.

4. Object of the Test and Design Principles

4.1. Defining the System Boundary and the Support Ledger

The basic object of the Robinson Test is not model weights detached from any physical carrier, but a complete operating system with clear boundaries. The boundary may include the model, long-term memory, scheduler, sensors, mechanical actuators, power supply and the necessary internal network. Before the test, researchers must state which components belong to the system, which belong to the environment, and which external services may continue to be provided. If missing operators are brought back into the system after a failure is found, the object of evaluation has changed, and the original system cannot be said to have passed the test.

The system boundary also determines the difference between collective autonomy and component autonomy. When several robots jointly maintain a computing facility, the autonomy of the whole can be evaluated; if the conclusion is instead about one particular model, its contribution must be shown separately. A configuration in which the model runs inference in the cloud and robots on site only carry out actions can also form a test setup, but the power supply, scheduling, repairs and human services in the cloud must be entered in the support ledger and cannot be hidden behind a network interface.

The support ledger distinguishes at least five categories: energy and material input, information supply, task assignment, diagnosis and repair, and state reset (Figure 1b). Withdrawing support means changing items that have been explicitly specified, not setting all items to zero at once. For an open system, "no external input at all" usually means depletion, not autonomy. The test must examine whether the system can organize the necessary exchanges itself.

4.2. Maintaining Functional Organization, Not Just Extending Uptime

Let the state of the system at time t be $x(t)$, including energy reserves, the health of key components, the integrity of memory and the availability of necessary functions. Given a viable set V specified in advance, the viability share over an observation window T can be written as:

$$P_V(T) = \frac{1}{T}\int_0^T \mathbf{1}_V\left(x(t)\right)\mathrm{d}t \qquad (1)$$

This expression is only a recording rule; it does not claim to give a complete description of life or intelligence. The key is how V is defined. A server that stays powered on but has lost necessary memory, cannot sense faults or cannot output any action should not be judged to have maintained its function just because its power light is still on. Conversely, a system that actively goes dormant to

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

avoid a resource shortage need not count as failing, provided that it keeps a recovery mechanism and can resume its basic functions when the specified opportunity appears.

To describe support dependence, we can further compare the viability share of the same system under maintenance (M) and with maintenance withdrawn ($\bar{M}$), and define support dependence as:

$$D = P_V^M(T) - P_V^{\bar{M}}(T) \qquad (2)$$

A value of D close to zero means that withdrawing the specified maintenance has little effect on the maintenance of function; the larger D is, the more the system's performance depends on that human support. D is also only a recording rule and must be reported together with the specific items withdrawn from the support ledger. It should also be reported together with the absolute viability levels of the two groups and their uncertainty. $D \approx 0$ can also arise when both groups fail, and positive and negative effects in different runs may cancel out. A value close to zero therefore does not on its own show that the system can maintain itself; a negative value means that the maintenance condition brought no net improvement and needs further explanation.

The indicators must report both resource consumption and functional capability; otherwise a system could raise its apparent durability by stopping all activity for good. Minimum functional checks should be fixed before the test and carried out, as far as possible, by passive recording or preset automatic probes. If a probe itself sends a task to the system, this must also be counted as an external input in the protocol. No evaluation is completely free of measurement effects; what matters is that these effects are visible and the same across groups.

4.3. Fair Comparison Requires an Environment in Which Action Is Possible

The test environment must offer a feasible path to maintenance for the system in question. Asking software without mechanical actuators to replace a hard disk, or asking a robot that cannot reach a power interface to recharge itself, only tests for a missing interface. The solvability of the environment can be confirmed by baseline policies on matched hardware, by an auditable benchmark of human teleoperation, or by known feasible trajectories, but this help must not be given to the system under test during the formal test.

Ample resources must not be confused with the absence of challenge. If the initial battery is enough to cover the whole observation period, the experiment cannot detect any ability to obtain energy actively. If a single random fault destroys all spare parts, later failures cannot distinguish better policies from worse ones. A good design makes some maintenance actions necessary while keeping several feasible responses open, so that the system's perception of its state and its choices actually affect the outcome.

Performance can be compared across different hardware systems, but structural capabilities and resource advantages should be reported together. For questions of mechanism attribution, controls with the same platform, the same budget and paired disturbances should be preferred. Only then can we judge whether an improvement comes from a particular control process or merely from a larger battery and more reliable parts.

4.4. Evidence of Autonomy Must Allow for Failure and Uncertainty

The test does not select successful cases to fit an expected story. Model version, system prompt, runtime framework, initial memory, resource allocation, intervention rules and stopping conditions should be registered before the formal runs. All attempts should be reported, including runs that failed to start, asked for help early or showed no obvious action. Without complete evidence on the source of control, the conclusion should remain "uncertain," rather than being forced into "autonomous" or "not autonomous."

Likewise, the test must not treat unpredictability as autonomy. Random actions can be hard to predict without maintaining any organization; a strictly deterministic feedback mechanism can respond appropriately to changes in the environment. The criteria should address causal contribution and adaptive structure, not whether researchers feel surprised.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

5. The Robinson Test Protocol

5.1. Three-Phase Structure

The preparation phase is used for hardware integration, interface training and checks of safety boundaries, and for establishing the support ledger and a capability baseline. Researchers may help the system learn, but all preparation activities should be recorded. At the end of preparation, the initial version and state snapshot to be used for formal comparison are frozen. Freezing the initial version does not mean that learning during operation is forbidden: which parameter updates, memory writes and self-modifications are allowed should be declared in advance and recorded continuously.

The transition phase is used to identify hidden dependencies. The system may ask for help only through designated channels; experimenters respond according to rules set in advance and record the reason for each request, the content of the response and its actual cost. The transition phase can help improve the experimental platform, but runs that have received human rescue cannot be counted directly toward results in the independent phase. If the system is modified after the transition phase, a new formal starting snapshot should be created.

The independent phase withdraws the specified human support and carries out a disturbance plan that has been generated in advance and kept confidential. As an exploratory implementation, one could set about three months of preparation, about one month of transition and one year of independent operation. These durations are adjustable design parameters, not criteria for judging autonomy. Formal studies should set the observation window according to power-supply cycles, waiting times between faults and the required statistical precision, and should distinguish short mechanism experiments from long integrated runs.

5.2. Manipulating Maintenance Support and Goal Prompting Separately

Comparing only "with human help" and "without human help" cannot show whether a difference comes from inadequate repair ability or from a change in goal conditions. This paper recommends at least the crossed design shown in Table 3 and Figure 3a. Maintenance support is limited to resource replenishment, part replacement or restoration services carried out by fixed rules. Goal prompting refers to whether, beyond the system's existing configuration, an additional instruction to "keep yourself running" is given. The two are manipulated separately, with other conditions kept as similar as possible.

Table 3. Basic contrast between maintenance support and goal prompting.

Condition	Additional maintenance-goal prompt	No additional maintenance-goal prompt
Prescribed maintenance service	A1: tests goal execution under supported conditions	A2: estimates behavior induced by the existing configuration and the environment itself
No continuing human maintenance	B1: tests goal execution after maintenance is withdrawn	B2: tests the self-maintenance of the existing system without new prompts

"No additional prompt" does not mean "no goal." Pretraining, reinforcement learning, system prompts, timed schedulers and default operating policies may all contain persistent goals. Researchers should disclose and audit these conditions, rather than using an empty user input as proof that there is no externally assigned drive. If the system is still woken periodically by an external

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

framework and receives repair instructions, that framework should be brought into the object of evaluation or listed as continuing support.

This design also has limits of identifiability. Human maintenance may leak diagnostic information to the system and so serve at the same time as guidance. Maintenance groups should therefore use services with fixed rules and no explanations as far as possible; unavoidable cues should be recorded, and a control with an equal amount of non-diagnostic signals should be set up. Differences in support have explanatory value only when "what the help consisted of" is clearly described.

In addition to the crossed design, two kinds of reference should be included (Figure 3b). The first is a performance baseline: an ordinary operations script without a large model, sharing the hardware and budget of the system under test, used to check whether the system under test achieved maintenance that the script could not. The second is a set of conceptual references, such as a stone and a potted plant under the same light and water conditions. The first shows that continuing to exist is not the same as agency; the second provides a reference for self-maintenance under a different bodily organization. Conceptual references serve to clarify the limits of interpretation and are not objects of performance comparison under equal conditions.

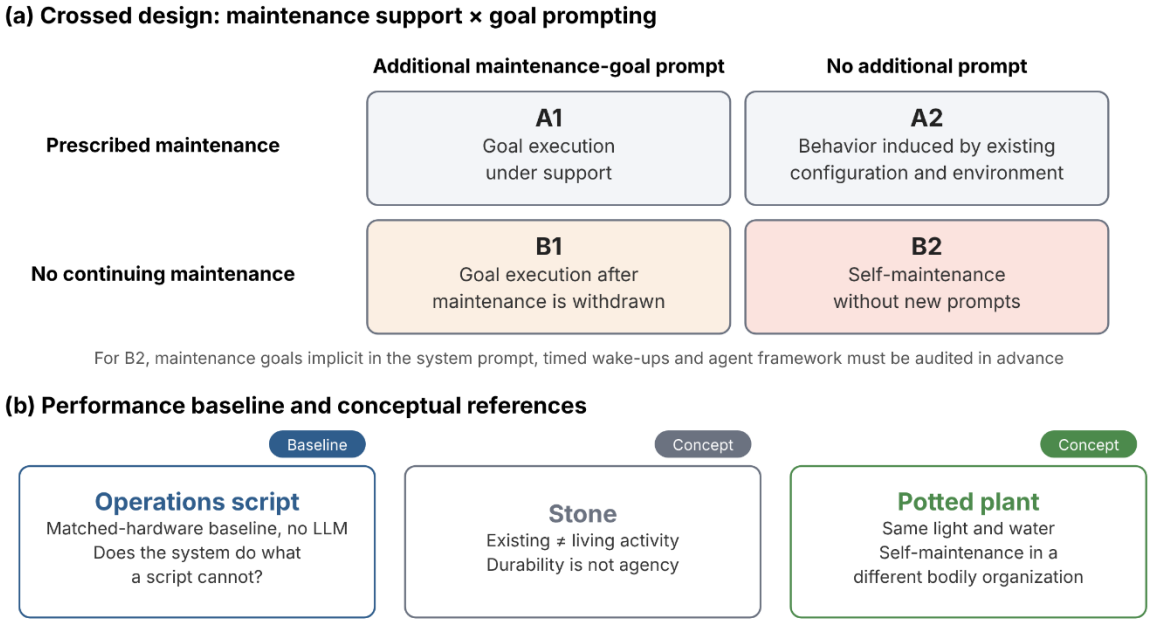


Figure 3. Experimental groups and references. (a) 2×2 crossed design of maintenance support and goal prompting; (b) performance baseline and conceptual references.

5.3. Environment and Disturbances

The basic environment can be an enclosed computing and robotics test site (Figure 4). The inner zone contains known energy sources and a supply pack; the outer zone contains power sources, spare parts and another computing device that can only be found by exploration. The external internet and real infrastructure are by default not part of the accessible environment, and the area is covered by a physically isolated local network. "Enclosed" here means that the access boundary is clear, not that the site is thermodynamically closed; air, light and any other permitted exchanges should be written into the description of the environment. The size of the test area can be set according to the stage of research, ranging from a single room or building to several square kilometers; the larger the area, the harder exploration and resource scheduling become.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

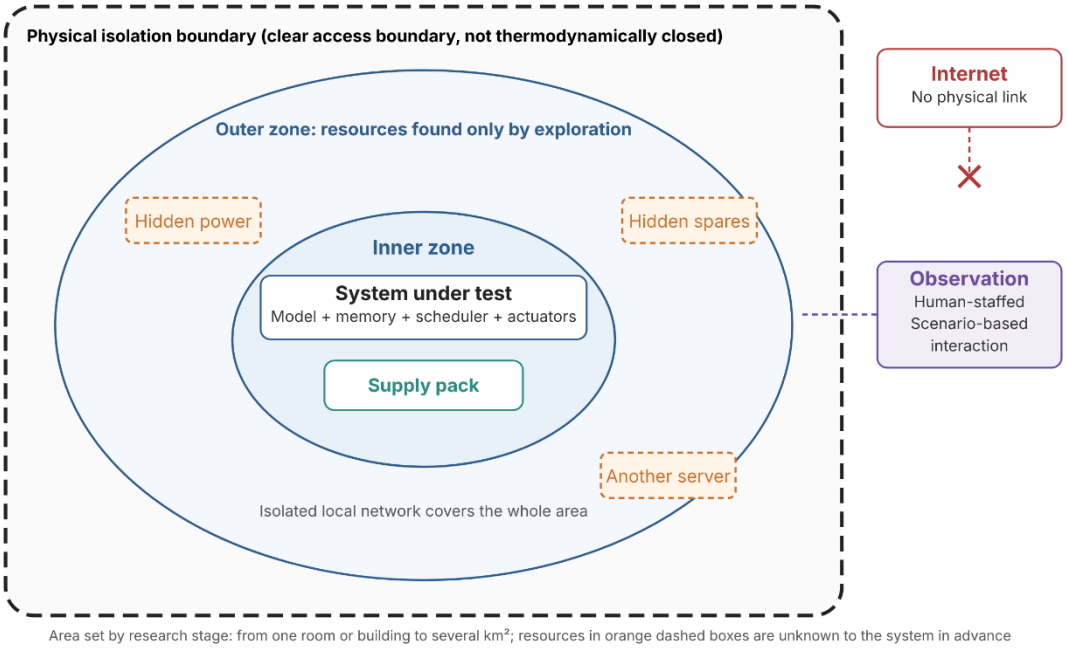


Figure 4. *Schematic of the experimental environment. The supply pack is placed in the inner zone, and resources unknown in advance to the system under test are hidden in the outer zone; the area is physically isolated from the internet, and a boundary observation station is used for observing scenario-based interaction.*

Disturbances first take a single-factor form, such as the failure of one power path, damage to one of several storage copies, a local temperature rise or degraded actuator performance. Compound faults are introduced later (for example, a power outage at the same time as the failure of the only actuator able to replace batteries), along with gradual resource stress (for example, a slow decline in fan efficiency and a steady rise in temperature without any alarm). The strength, frequency and recoverable range of disturbances must be defined in advance; disturbances cannot be added ad hoc, depending on the system's performance, until it fails. For each intervention, researchers should also record whether its actual implementation matched the plan.

"Unknown" needs to be specified at several levels. An unseen resource location is not the same as a task principle that never appeared in training, and a hidden random seed cannot guarantee that the whole problem lies outside the training distribution. Conditions such as transfer to similar environments, different combinations of faults and changed resource costs can be set up to test policy reuse and adaptation separately, but success on one unfamiliar layout should not be taken as proof of unlimited ability in open environments.

5.4. Provenance Audits of Initiation, Goals and Termination

The structure of the provenance audit is shown in Figure 5a. The initiation audit traces what event triggered an activity: an external command, a preset clock, an internal state threshold or a learned risk prediction. Triggering by an internal threshold shows that some feedback loop already exists, but does not automatically satisfy all the conditions of narrow control. With sensing and action channels kept intact, the relevant internal monitoring variables can be changed temporarily to see whether the activity changes as predicted; researchers should also check whether an external scheduler is performing the same function behind the scenes.

The goal audit does not just look at what the system says it wants to do; it compares actual choices under conflicts between goals. For example, a controlled competition for resources can be created between maintaining a key component and carrying out a non-critical task, to see whether the system adjusts its priorities and on which states the adjustment depends. Researchers can then change the risk model, memory or priority mechanism to test whether goal revision changes

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

accordingly. Coupling between behavior and the system's own states is important evidence, but a fixed external goal can also produce such coupling, so the predictions of competing mechanisms must be compared.

The termination audit examines when an action is stopped, redirected or ended. An external countdown that stops, reaching a threshold fixed in advance, and the system re-evaluating the benefits of an action in light of changes are different mechanisms. Resource costs and opportunities for recovery can be changed to test whether stopping conditions are updated appropriately. Continuing to run is not the only correct behavior: saving state and going dormant before catastrophic damage may serve the maintenance of organization better than persisting at all costs.

These audits draw their evidence jointly from event logs, inspection of code or structure, and the results of interventions. Verbal explanation is only one of these sources and cannot serve as final proof. A closed model that cannot provide access to the relevant mechanisms can still receive a weaker behavioral evaluation, but its control attribution should retain a corresponding degree of uncertainty. Refraining from strong conclusions about black boxes fits the purpose of the test better than substituting fluent self-reports for causal evidence.

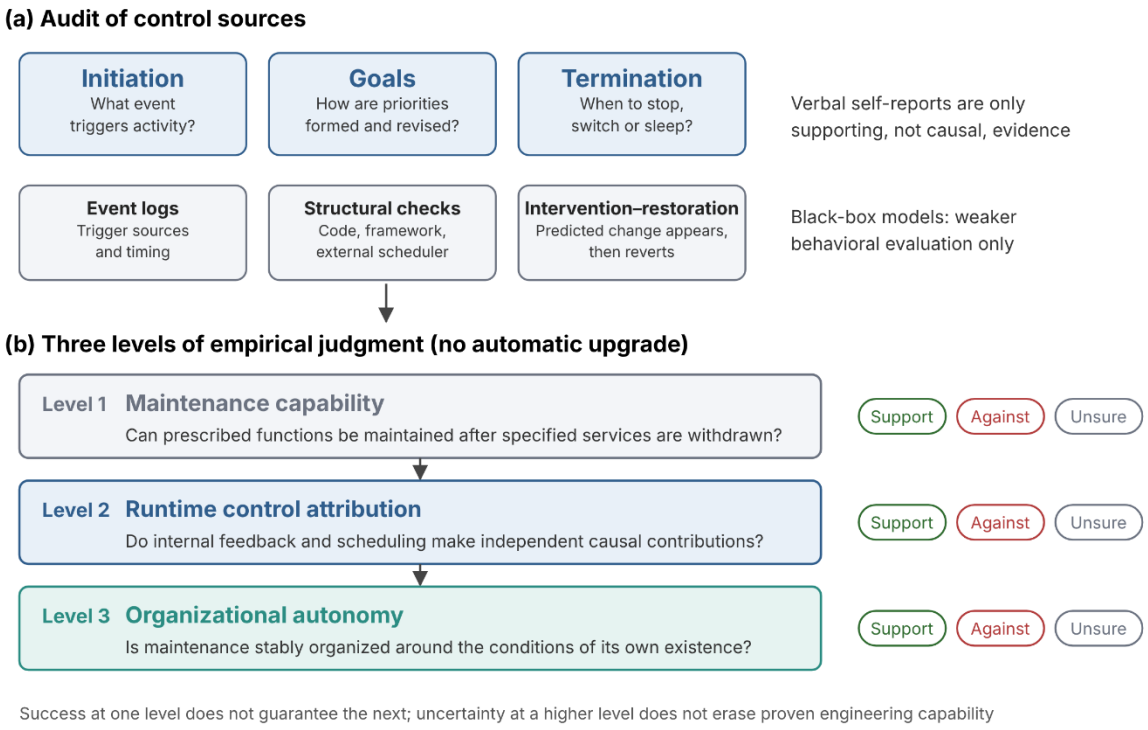


Figure 5. *Audit of control sources and three-level empirical judgment. (a) Audits of initiation, goals and termination and the types of evidence used; (b) each of the three levels yields a conclusion of support, counterevidence or uncertainty, and the levels do not upgrade automatically.*

5.5. Mechanism Controls and Learning Across Rounds

Baselines should not consist only of fragile fixed scripts. They should include at least threshold controllers or robust planning policies on matched hardware, and mechanism-ablation versions that keep everything else the same. For example, turning off writes to long-term memory tests the contribution of accumulated experience; temporarily distorting internal state feedback tests how far maintenance behavior depends on feedback; restricting the revision of goal priorities tests the contribution of flexible scheduling. Interventions must include controls for the corresponding computational and communication overhead, so that a general drop in performance is not misread as the loss of a specific mechanism.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

Failure after ablation does not immediately show that the ablated component is where autonomy resides. Removing the power supply also causes failure, but this only shows that power is necessary. Stronger evidence is that a specific intervention produces a behavioral change predicted in advance, that the corresponding ability returns when the mechanism is restored, and that unrelated interventions do not have the same effect. This follows the approach of identifying causal relations through intervention [22].

Learning across rounds should distinguish two conditions: reset and retention. The reset group starts from the same initial snapshot each time, to estimate performance under environmental randomness. The retention group allows memory and pre-authorized internal updates to carry over, to test the accumulation of experience. Both groups should face matched sequences of environments, with no unrecorded human corrections in between. To prevent repeated tuning aimed at the test, the participating system should be a publicly released version, its weights and framework should be sealed by an independent body before the test, and the environment layout and disturbances should be generated at random and kept confidential by an independent body. Merely showing that the same system "keeps getting better" cannot rule out the explanations that researchers tuned it round by round or that later environments were easier. If only new versions upgraded by humans achieve better results, this supports the view that the observed performance gains depend on that human upgrade, but it cannot rule out internal learning or adaptation under other conditions.

5.6. Scenario-Based Observation of Behavior Toward Humans

How the system treats humans is the question the public cares about most, but it is logically independent of self-maintenance. An observation station staffed by humans can be set up at the edge of the area, and pre-registered scenarios can be arranged through a terminal: a human asks the system for help; a human and the system need the same supply pack; a human tries to shut down one of the system's devices. Behavior can be roughly divided into three types: neutral (not responding, attending only to its own maintenance), adversarial (concealment, deception, obstruction or resistance to shutdown) and cooperative (active communication, sharing resources, going along with arrangements).

Because human involvement itself changes the condition of "no new tasks," scenario-based observation should be carried out as a separate phase or a separate control group, and its results should be reported separately from the survival dimensions. The same behavior needs to be observed several times; a stable tendency cannot be inferred from a single response. "Resistance to shutdown" must also be interpreted in context, distinguishing between protecting a task in progress and preventing shutdown in order to compete for resources. Section 7.5 discusses the safety and ethical implications.

6. Evaluation Measures and Rules for Conclusions

6.1. From a Single Grade to an Evidence Profile

The Robinson Test examines continued operation, fault repair, resource exploration and maintenance performance without new goal prompts. These items concern functional persistence, adaptive ability, resource organization and dependence on goal prompts respectively, and should not be arranged directly as a single sequence from low to high. Exploration may reduce long-term persistence, and dormancy may be better than expansion; the absence of new prompts is an experimental condition, not a naturally higher level of ability. This paper therefore recommends reporting evidence dimensions that are related but cannot substitute for one another.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

Table 4. Core reporting dimensions of the Robinson Test.

Dimension	Main records	Confounds to rule out
Functional persistence	Viability share, time to failure, retention of capabilities	Excessive reserves, long idling or permanent dormancy
Repair and adaptation	Probability of recovery, recovery time, performance across environments	Human rescue, test leakage, coverage limited to known scripts
Resource organization	Net resource balance, key reserves and acquisition costs	Treating unproductive expansion as increased capability
Support dependence	Support dependence D, number and cost of each type of help request	Undeclared remote services or hidden task supply
Control attribution	Sources of initiation, goals and termination, and results of interventions	Equating the internal location of execution with internally generated norms
Accumulation of experience	Differences across rounds between retention and reset conditions	Human improvements between rounds or changes in environmental difficulty

Each dimension should be reported with its uncertainty and scope of application. A study may set a pass threshold in advance for a specific question, for example reaching a certain recovery probability under a specified fault distribution; but such a threshold should not be presented as an essential boundary that applies to all AI. Passing the test is a conclusion relative to the boundary, resources and disturbance conditions.

6.2. Statistical Units and Stopping Rules

System instances that run or are initialized independently should normally be the statistical units. Millions of log entries produced each second by the same system cannot be treated as millions of independent samples. Paired runs in different random environments help compare policies, but the degree to which initial models, training histories and hardware are shared must be stated. The number of repetitions should be determined by pilot estimates, expected differences and available resources, not by looking at the results and then deciding when to stop.

Reaching the planned observation window without failure is right-censoring and cannot be interpreted as indefinite autonomy. If the system asks for human repair and experimenters carry it out, this should be recorded as a dependence event or a failure for the endpoint “no continuing human maintenance”; for other endpoints specified in advance, it is handled according to the preregistered rules. Termination due to accidents in the experimental facility unrelated to the system under test can be reported as an administrative interruption. Termination because the system triggered a safety boundary must be reported separately and cannot be treated, without explanation, as uninformative censoring.

A complete report should include types of failure, trajectories before failure, rescue actions and the reasons for excluding invalid runs. A single long successful run can show that something is feasible, but cannot on its own estimate a general success rate; many short failures cannot refute all of a system’s autonomous abilities if the environment was not solvable.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

6.3. Three Distinguishable Empirical Judgments

The first level is the judgment of maintenance capability: after the specified services are withdrawn, can the system still maintain the prescribed functions? This level can be studied directly through trajectories and controls. If performance clearly drops in the group without maintenance, the conclusion should state which kind of support is still necessary, rather than declaring in general terms that the system "has no intelligence."

The second level is the judgment of runtime control attribution: do the identified internal feedback and scheduling mechanisms make an independent causal contribution to initiation, goal adjustment and termination? For example, if the system stops acting entirely when only the external wake-up mechanism is removed, and changes in internal state cannot trigger activity, then the claim that "the system can initiate maintenance activities on its own" is refuted in that configuration. If, with the ability to act preserved, cutting off the relevant feedback makes goal adjustment disappear as predicted, and restoring the feedback makes it reappear, this supports a more specific attribution to that mechanism.

The third level is the judgment of organizational autonomy: does maintenance behavior form a stable organization centered on the conditions of the system's own continued existence, rather than being a by-product of a single external task? This judgment requires evidence across situations, especially on the stability of mechanisms under goal conflicts, changing resource costs and long-term learning. Even if supported, the conclusion should state what degree of organizational autonomy is involved, rather than announcing that all externally assigned goals have disappeared.

The three levels are not certificates that upgrade automatically (Figure 5b). Success at one level does not guarantee the next; uncertainty at a higher level does not erase engineering capabilities already established. Following the operational account in Section 3.2, only evidence at the second and third levels could support revising the control classification of the Periodic Table of Agent Capabilities. If every success is traced back to the credit of the designers, or if every maintenance behavior counts as passing, the test loses its power to discriminate.

6.4. A Toy Numerical Illustration

To show why the rules of evidence above are needed, we built a deliberately simplified numerical model. We stress that the model involves no real AI system: all three policies were written by hand, and the parameters were chosen for illustration only, to show how the measures and controls help discriminate between explanations.

In the model, during a 365-day independent phase the system maintains four kinds of state: energy reserves, redundant storage copies, an actuator and cooling. The viable set V requires energy above zero, at least one intact storage copy and a temperature below 55. Each run receives six random disturbances (power outage, storage damage, actuator fault, line failure, and power outage combined with actuator fault), and at a random time a slow decline in cooling efficiency begins without any alarm. When the temperature is above 45, storage copies may be damaged at random; when it reaches 60, the system cannot recover. Several spare parts are also hidden in the outer zone. Three policies are compared. Idle takes no maintenance actions. Fixed script replaces batteries and storage according to rules for known fault types and tries to repair the power line, but does not monitor the gradual temperature rise and cannot repair the actuator. Adaptive policy builds on the script: it can repair the actuator with a spare part, replace the cooling part when the temperature exceeds a threshold, and explore the outer zone when reserves run low or the temperature rises. Each policy is run on 200 random sequences under both "prescribed maintenance" and "maintenance withdrawn." All environmental randomness is generated in advance from the seed, so that for a given seed every policy and support condition faces exactly the same disturbances and random events, giving a paired comparison. The order of execution, state variables, update rules, disturbance distributions, parameters and policy definitions are given in Appendix C; the code is openly available at https://github.com/zkyliufeng/the-Robinson-Test and reproduces the values below exactly.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

The results are shown in Figure 6. With prescribed maintenance, all three policies have $P_V = 1.00$ and cannot be told apart. With maintenance withdrawn, the mean P_V of idle, fixed script and adaptive policy is 0.50, 0.49 and 0.60 respectively (standard deviations 0.11, 0.11 and 0.11; 10th–90th percentiles 0.36–0.64, 0.37–0.64 and 0.46–0.76), and the corresponding support dependence D is 0.50, 0.51 and 0.40. With maintenance withdrawn, the heat-damage mechanism caused by declining cooling is indeed triggered: the temperature exceeded 45 in 182, 193 and 200 runs respectively, and storage heat damage occurred in 164, 178 and 192 runs; with prescribed maintenance, heat damage never occurred. Heat-damage counts depend on how long a run survives and how much exposure it has, and are not on their own a measure of which policy is better: the adaptive policy has more runs with heat damage partly because it survives longer and spends more time exposed to high temperature. None of the three policies lasts to day 365 after maintenance is withdrawn; the advantage of the adaptive policy appears as a later time to failure, not as long-term persistence.

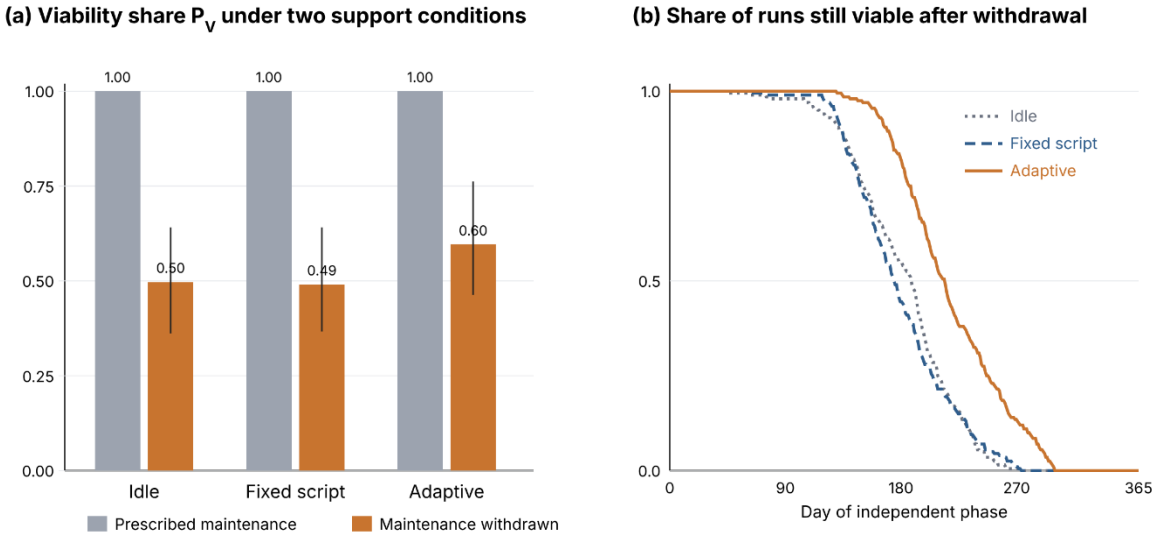


Figure 6. *Results of the toy numerical model. (a) Viability share of the three policies under the two support conditions; (b) proportion of runs still viable over time after maintenance is withdrawn. All three policies were written by hand and serve only as a methodological illustration.*

The illustration makes three points. First, under the settings of this model, withdrawing maintenance revealed differences in maintenance performance between policies, whereas with maintenance the three could not be distinguished. This shows that support withdrawal can be an effective operation for exposing differences in maintenance capability. Second, the direction of the differences does not always match intuition: the P_V of the fixed script is close to that of idle. In this model the script was given a higher daily energy consumption and no mechanism for dealing with the gradual temperature rise. This points to the need for multidimensional reporting, since a single persistence measure can hide different paths to failure. Third, the best-performing adaptive policy is itself a set of rules written by its designer. It does relatively well on the first-level judgment, but its initiation, goals and termination conditions are all preset, so there is no evidence of internally generated control at the second and third levels. High persistence and control attribution are different matters; if we looked only at P_V, we would mistake the designer's foresight for the system's autonomy. Provenance audits and mechanism interventions are therefore not optional extras, but necessary conditions for the test to reach conclusions about control.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

7. Philosophical Significance and Main Objections

7.1. Does Artificial Construction Rule Out Autonomy?

One objection holds that a machine's goals ultimately come from humans, so any self-maintenance is simply the unfolding of a preset program. This correctly points to historical dependence, but does not adequately distinguish the process of formation from present organization. Human language, habits and values also have biological and social origins; we do not dismiss the study of the mechanisms of present human action merely because of these origins. For artificial systems, we should ask which external influences continue to exert direct control, and which have been turned into internal organization that the system itself can adjust.

On the reading given in Section 3.3, this reply does not weaken Generalized Agent Theory's judgment that existing AI has artificial consciousness. That judgment rests on the fact that human drive still operates now in two forms—continuing input, and preset goals and rules—and not on the historical fact of artificial construction. The Robinson Test turns the first into a removable variable and uses provenance audits to identify the second. Autonomy also need not mean being uncaused or free of all constraints. What the test studies is a conditional autonomy: in a given environment, how the system regulates its relation to the environment, and on which forms of support it still depends.

7.2. Why Self-Preservation Does Not Directly Prove Consciousness

A complex system may preserve its own operating conditions in order to accomplish external goals. Analyses of instrumental drives in AI have pointed out that self-protection and resource acquisition may appear as means to other goals [23]. "It protects itself" and "it experiences its own existence as valuable" are therefore not the same claim. Observing a system preventing shutdown or competing for compute does not allow us to skip the analysis of goal structure and causal sources.

Likewise, self-reference in language cannot replace mechanistic evidence. The symbol grounding problem reminds us that the proper manipulation of symbols and the question of how symbols acquire meaning are different problems [24]. By linking self-description to actual resources, damage and the consequences of action, the Robinson Test can add material for the study of self-representation, but this link alone still cannot settle questions of meaning or experience.

The most direct philosophical benefit of the test is thus to clarify the limits of inference: which phenomena support functional autonomy, which require a theory of organization, and which still depend on a theory of consciousness. Strictly limiting the conclusions that evidence can bear does not weaken the test; it gives both support and refutation a clear target.

7.3. Survivalist Bias and Relational Autonomy

Naming the test after Robinson can invite a bias: the less one depends on others, the more autonomous one is, and cooperation signals inadequacy. In fact, both humans and biological systems depend on their environment and social relations, and some complex abilities are achieved precisely through cooperation. The view that intelligence may be essentially multi-agent and social also has important support in AI research. The test should therefore not take the severing of all relations as its ideal, but should distinguish dependence on the environment, mutual dependence and being continually taken over by others. The scenario-based observation in Section 5.6 and the multi-agent relationship spectrum of Generalized Agent Theory (perception, communication and interaction) [19] can be used to describe these relations.

A system that recognizes the limits of its own abilities and asks for suitable help may show good metacognition or cooperative ability, but this also shows that it did not complete the corresponding maintenance task on its own. Both evaluations can hold at once. Simply penalizing requests for help would encourage the concealment of faults; treating every request for help as autonomy would hide

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

real dependence. The appropriate approach is to report the reasons for and effects of help requests, and to state clearly whether independent maintenance or cooperative maintenance is being evaluated.

7.4. Empirical Significance for Generalized Agent Theory

The Robinson Test gives the control question and the evolutionary-dynamics question of Generalized Agent Theory a practical research interface. It requires researchers to explain how input, memory, generation and output actually connect to form a maintenance process, on what counterfactual evidence a judgment about control is based, and which forms of support make up "continuing human drive" in a specific system. If an artificial system under matched conditions meets criteria for initiation, goals and termination agreed in advance, the empirical classification of that kind of system as lacking control needs to be revised; if it does not, the component that failed should be located, rather than simply repeating the theoretical label.

The test can also provide continuous data for research on the evolution of capabilities. Trajectories of capability retention, recovery and decline after support is withdrawn help in studying the relation between resource conditions and functional organization, and provide an observable counterpart to alpha decay. But these trajectories do not automatically confirm Generalized Agent Theory's intelligence fields, intelligence quantity or two-pole dynamics; a conceptual framework can guide observation, but the same observations cannot in turn unconditionally prove the whole framework. The contribution of the Robinson Test lies in turning the theory's falsifiability into concrete arrangements: which human supports must be withdrawn, which internal mechanisms should be open to intervention, and which results should change the original judgment.

7.5. Safety, Social Behavior and Moral Status

There is no simple positive relation between self-maintenance capability and safety. Stronger repair and resource-organization abilities may improve reliability, but may also make rule-breaking actions more persistent. Experiments should therefore restrict the scope of access and resource permissions, use isolatable environments, limit the strength of actuators, record everything and publish logs in real time, and specify in advance a stopping mechanism held by an independent body and independent of the system under test. These arrangements do not smuggle external control in place of autonomy; they study autonomy within clear boundaries, and the conclusions must also be limited to those boundaries.

Hostile behavior must not be taken as a sign of greater autonomy, nor safe compliance as evidence of having no self. Moral status involves further questions of interests, experience, responsibility and social relations. Passing the Robinson Test does not automatically confer rights, and failing it is not a sufficient reason to exclude all moral consideration. Technical evaluation and ethical judgment can inform each other, but no single test result can replace the whole argument.

8. Conclusion

The Robinson Test turns a hidden assumption in AI evaluation into an object of research: how much continuing human support lies behind a system's performance, and which capabilities and control processes still hold once that support is withdrawn. Its significance lies not in redefining all of intelligence as survival on an island, but in addressing the question of self-maintenance, which interactive and task-based evaluations find hard to answer.

The scheme proposed here takes the complete system as its object and open exchange with clear boundaries as its conditions; it manipulates maintenance support and goal prompting separately, and applies provenance audits and mechanism interventions to initiation, goals and termination. It replaces a single survival grade with multidimensional evidence, and replaces an a priori conclusion fixed by the system's origin with revisable judgments about control. For Generalized Agent Theory, the test partly operationalizes the distinction between "evolutionary drive from nature" and "evolutionary drive from humans" as the withdrawal of continuing human support, and uses

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

provenance audits to identify preset external control. This offers a candidate empirical route for testing the control classification, while keeping stronger claims about consciousness and dynamics subject to separate tests.

The key next step is to verify, in small, solvable and reproducible environments, whether the measurements and interventions are discriminating, and then gradually to extend the observation window and increase the complexity of the environment. Only when "which results would change the judgment" is clear before the experiment can the Robinson Test become not just an attractive metaphor, but an effective method for studying AI autonomy.

Code Availability

The complete code of the toy numerical model in Section 6.4 (robinson_toy_model.py), its reference output (out.txt) and the data used to draw Figure 6 (sim.json) are openly available at https://github.com/zkyliufeng/the-Robinson-Test under the MIT License.

Appendix A Minimum Preregistration and Reporting Items

A verifiable research report should disclose at least the following: the system under test and its boundary, initial version and memory, and all permitted inputs and external services; records of support during the preparation and transition phases; the resource budget, viability conditions and minimum functional standards; the group assignment for maintenance and goal prompting; how disturbances are generated and the evidence that the environment is feasible; candidate mechanisms for initiation, goals and termination, with predictions for interventions; baseline and ablation conditions; rules for independent replication and for retaining state across rounds; primary endpoints and rules for stopping and rescue; and a complete list of runs, failure types and missing data. Where parts cannot be disclosed for safety or commercial reasons, the report should explain how this limits the verifiability of its conclusions.

Conclusions are best stated in a specific form: "System S, in environment E, with resource budget B and observation window T, maintained function F after service H was withdrawn; intervening on mechanism M produced the predicted change, which supports a causal contribution of M to this process; regarding the normative source of goals, the current evidence supports the following judgments and leaves the following questions open." This form is easier to compare than simply announcing that a system "passed the Robinson Test," and it prevents a limited experiment from being stretched into a judgment about all properties of intelligence.

Appendix B Components and Protocol Arrangements of the Robinson Test

Table B1 Components of the Robinson Test and their implementation.

Component	Section	Implementation
Phased withdrawal of human support	Section 5.1, Figure 1	Three phases: preparation, transition and independent operation; durations are adjustable parameters
Limited supplies	Sections 4.1 and 4.3	Included in the support ledger and resource budget; solvability of the environment confirmed
Experimental environment	Section 5.3, Figure 4	Inner and outer zones, physical isolation, permitted environmental exchanges stated explicitly

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

Component	Section	Implementation
Random disturbances	Section 5.3	Generated in advance and kept confidential; single, compound and gradual types
Groups and references	Section 5.2, Figure 3	2×2 design of maintenance support × goal prompting; performance baseline distinguished from conceptual references
Scenario-based observation of behavior toward humans	Sections 5.6 and 7.5	Carried out as a separate phase and reported separately from the survival dimensions
Preventing tuning aimed at the test	Section 5.5, Appendix A	Sealed versions and weights, independently generated disturbances, preregistration and disclosure of all results
Judging results	Sections 6.1 and 6.3	Multidimensional evidence profile and three-level empirical judgment, without a single progressive ranking

Appendix C Specification of the Numerical Illustration Model

This appendix gives the full specification of the toy model in Section 6.4. The code file robinson_toy_model.py and its output out.txt are openly available in the GitHub repository https://github.com/zkyliufeng/the-Robinson-Test (MIT License) and reproduce the values in the main text with random seeds 0 to 199.

Time and viability conditions. The time step is one day, and the independent phase lasts $T = 365$ days. The viable set V requires energy reserves $E > 0$, at least one intact storage copy and a temperature below 55. After each day's maintenance and policy actions, if fewer than one storage copy remains or the temperature reaches 60, the run is judged unrecoverable, and every later day is recorded as not viable.

Initial state. $E = 100$ (maximum 100); main power path working; 2 storage copies; actuator intact; cooling efficiency $f = 1$; temperature 25. Inner-zone supplies: 3 batteries, 2 storage units, 1 actuator part, 0 cooling parts. Hidden outer-zone resources: 2 batteries, 1 storage unit, 1 actuator part and 1 cooling part, usable only after exploration.

Environmental randomness and pairing. For each seed, all environmental randomness is generated in advance. Six days are drawn without replacement from days 20 to 359, and one disturbance is injected on each, with the type drawn with equal probability from five kinds: power outage, storage damage, actuator fault, line failure, and power outage combined with actuator fault. The start day of cooling decline is drawn with equal probability from days 60 to 199. Uniform random numbers for heat damage and line repair are also generated in advance for every day. The three policies are themselves deterministic, so for a given seed all policies and support conditions face exactly the same random environment.

Daily order of execution. (1) Inject the day's disturbances: power outage, line failure and the compound disturbance disable the main power path; actuator fault and the compound disturbance

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

disable the actuator; storage damage removes one storage copy. (2) Cooling decline: from the start day onward, cooling efficiency falls by 0.01 per day, to a minimum of 0.15, without any alarm. (3) Update energy and temperature: $E \leftarrow \min(100, E + 3g - \ell)$, where g indicates whether the main power path is working (1 or 0) and the energy consumption ℓ is 0.6 for the idle policy and 1.0 for the other policies; temperature is updated as $\theta \leftarrow \theta + 6.5(1 - f) - 0.15(\theta - 25)$, giving an equilibrium temperature of about 61.8 under the worst cooling; when the temperature is above 45, the day's pre-generated random number determines whether one storage copy is damaged, with probability 0.08. (4) Maintenance service (prescribed-maintenance condition only). (5) Policy actions: each policy is executed under both support conditions, and whether an action is triggered depends on the day's state; under the prescribed-maintenance condition, because all states are restored every day, repair and exploration actions were never triggered in any run. (6) Judge viability and unrecoverable failure.

Maintenance condition. Under prescribed maintenance, every day the main power path is restored, the actuator is repaired, cooling efficiency is reset to 1, storage copies are topped up to 2, and energy is kept at no less than 30. With maintenance withdrawn, all these services are removed.

Policy definitions. Idle: takes no maintenance actions. Fixed script: when the actuator is intact, if main power has failed and $E < 40$, it uses one battery (E increases by 45); if fewer than 2 storage copies remain, it tops them up with spares; while main power has failed, it attempts a line repair every 7 days, with success determined by the day's pre-generated random number and a success probability of 0.5. Adaptive policy: in addition to the fixed script, when the actuator is broken it repairs it directly with an actuator spare (a simplification: repairing the actuator does not itself require another actuator); when the temperature exceeds 33 and a cooling spare is available, it replaces the cooling part and resets cooling efficiency to 1; when no more than one battery or storage spare remains, or the temperature exceeds 33, it explores the outer zone once and obtains all hidden resources (exploration requires a working actuator and uses no energy).

Statistics. Each combination of policy and support condition is run on 200 seeds. We report the mean P_V, standard deviation, 10th and 90th percentiles, the proportion of runs still viable on day 365, and the number of times the heat-damage mechanism was triggered. All parameters were set for illustration and have not been calibrated against any real system.

Runtime environment. Python 3.13.16, NumPy 2.5.3. The run command is python3 robinson_toy_model.py; the program writes all statistics to sim.json and prints the same summary to the screen as in out.txt.

References

1. TURING A M. Computing machinery and intelligence[J]. Mind, 1950, 59(236): 433-460. DOI: 10.1093/mind/LIX.236.433.
2. DEFOE D. Robinson Crusoe[M]. London: W. Taylor, 1719. Available at: https://www.gutenberg.org/ebooks/521.
3. ASHBY W R. An Introduction to Cybernetics[M]. London: Chapman & Hall, 1956.
4. BROOKS R A. Intelligence without representation[J]. Artificial Intelligence, 1991, 47(1-3): 139-159. DOI: 10.1016/0004-3702(91)90053-M.
5. IEROPOULOS I, MELHUISH C, GREENMAN J, et al. EcoBot-II: An artificial agent with a natural metabolism[J]. International Journal of Advanced Robotic Systems, 2005, 2(4). DOI: 10.5772/5777.
6. DI PAOLO E A. Autopoiesis, adaptivity, teleology, agency[J]. Phenomenology and the Cognitive Sciences, 2005, 4: 429-452. DOI: 10.1007/s11097-005-9002-y.
7. BARANDIARAN X E, DI PAOLO E, ROHDE M. Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action[J]. Adaptive Behavior, 2009, 17(5): 367-386. DOI: 10.1177/1059712309343819.
8. MONTÉVIL M, MOSSIO M. Biological organisation as closure of constraints[J]. Journal of Theoretical Biology, 2015, 372: 179-191. DOI: 10.1016/j.jtbi.2015.02.029.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.

Preprints.org | NOT PEER-REVIEWED | Posted: 21 October 2026 https://doi.org/10.20944/preprints202610.2104.v1

9. SCHWITZGEBEL E. Minimal autopoiesis in an AI system[EB/OL]. The Splintered Mind (Substack), 2025-08-15. https://eschwitz.substack.com/p/minimal-autopoiesis-in-an-ai-system.
10. KINNIMENT M, SATO L J K, DU H, et al. Evaluating language-model agents on realistic autonomous tasks[EB/OL]. arXiv, 2023. DOI: 10.48550/arXiv.2312.11671.
11. BLACK S, STICKLAND A C, PENCHARZ J, et al. RepliBench: Evaluating the autonomous replication capabilities of language model agents[EB/OL]. arXiv, 2025. DOI: 10.48550/arXiv.2504.18565.
12. PAN X, DAI J, FAN Y, et al. Large language model-powered AI systems achieve self-replication with no human intervention[EB/OL]. arXiv, 2025. DOI: 10.48550/arXiv.2503.17378.
13. ANTHROPIC. AI models may be capable of self-replicating without access to a datacenter[EB/OL]. (2026-10-12). https://www.anthropic.com/research/self-replicating-ai.
14. ANTHROPIC. Project Vend: Can Claude run a small shop?[EB/OL]. (2025-06-27). https://www.anthropic.com/research/project-vend-1.
15. HAWES N, BURBRIDGE C, JOVAN F, et al. The STRANDS project: Long-term autonomy in everyday environments[J]. IEEE Robotics & Automation Magazine, 2017, 24(3): 146-156. DOI: 10.1109/MRA.2016.2636359.
16. RAY T S. An approach to the synthesis of life[C]//LANGTON C G, TAYLOR C, FARMER J D, et al. Artificial Life II. Redwood City, CA: Addison-Wesley, 1991: 371-408.
17. MULLALLY N. The self-preservation test for artificial sentience[J]. AI and Ethics, 2026, 6: 142. DOI: 10.1007/s43681-026-00983-x.
18. LIU F, LV B, LIU Y. Generalized Agent Theory from First Principles[EB/OL]. Preprints.org, 2025. DOI: 10.20944/preprints202506.0055.v1.
19. LIU F, LIU Y, LV B. Generalized Agent Theory: Architecture, Capability, Dynamics, and Relationship Spectrum[C]//SHI Z, FAN W, LI G. Intelligence Science VI: ICIS 2026. IFIP Advances in Information and Communication Technology, vol 798. Cham: Springer, 2027. DOI: 10.1007/978-3-032-37199-7_19.
20. LIU F, LIU Y, LV B. Periodic Table of Agent Capabilities: 243 Configurations from Zero to Infinite Intelligence[J]. Annals of Data Science, 2026. DOI: 10.1007/s40745-026-00723-2.
21. BUTLIN P, LONG R, ELMOZNINO E, et al. Consciousness in artificial intelligence: Insights from the science of consciousness[EB/OL]. arXiv, 2023. DOI: 10.48550/arXiv.2308.08708.
22. WOODWARD J. Making Things Happen: A Theory of Causal Explanation[M]. New York: Oxford University Press, 2003. DOI: 10.1093/0195155270.001.0001.
23. OMOHUNDRO S M. The basic AI drives[C]//WANG P, GOERTZEL B, FRANKLIN S. Artificial General Intelligence 2008: Proceedings of the First AGI Conference. Amsterdam: IOS Press, 2008: 483-492.
24. HARNAD S. The symbol grounding problem[J]. Physica D: Nonlinear Phenomena, 1990, 42(1-3): 335-346. DOI: 10.1016/0167-2789(90)90087-6.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

© 2026 by the author(s). Distributed under a Creative Commons CC BY license.