



OPEN ACCESS

EDITED BY

Sheri Marina Markose,
University of Essex, United Kingdom

REVIEWED BY

José Antonio Cervantes,
University of Guadalajara, Mexico
Joshua Ljubo Bensemann,
The University of Auckland, New Zealand

*CORRESPONDENCE

Robert Johansson,
✉ robert.johansson@psychology.su.se

RECEIVED 29 May 2024

ACCEPTED 30 July 2024

PUBLISHED 14 August 2024

CITATION

Johansson R (2024) Machine Psychology: integrating operant conditioning with the non-axiomatic reasoning system for advancing artificial general intelligence research.

Front. Robot. AI 11:1440631.

doi: 10.3389/frobt.2024.1440631

COPYRIGHT

© 2024 Johansson. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Machine Psychology: integrating operant conditioning with the non-axiomatic reasoning system for advancing artificial general intelligence research

Robert Johansson^{1,2*}

¹Department of Psychology, Stockholm University, Stockholm, Sweden, ²Department of Computer and Information Science, Linköping University, Linköping, Sweden

This paper presents an interdisciplinary framework, Machine Psychology, which integrates principles from operant learning psychology with a particular Artificial Intelligence model, the Non-Axiomatic Reasoning System (NARS), to advance Artificial General Intelligence (AGI) research. Central to this framework is the assumption that adaptation is fundamental to both biological and artificial intelligence, and can be understood using operant conditioning principles. The study evaluates this approach through three operant learning tasks using OpenNARS for Applications (ONA): simple discrimination, changing contingencies, and conditional discrimination tasks. In the simple discrimination task, NARS demonstrated rapid learning, achieving 100% correct responses during training and testing phases. The changing contingencies task illustrated NARS's adaptability, as it successfully adjusted its behavior when task conditions were reversed. In the conditional discrimination task, NARS managed complex learning scenarios, achieving high accuracy by forming and utilizing complex hypotheses based on conditional cues. These results validate the use of operant conditioning as a framework for developing adaptive AGI systems. NARS's ability to function under conditions of insufficient knowledge and resources, combined with its sensorimotor reasoning capabilities, positions it as a robust model for AGI. The Machine Psychology framework, by implementing aspects of natural intelligence such as continuous learning and goal-driven behavior, provides a scalable and flexible approach for real-world applications. Future research should explore using enhanced NARS systems, more advanced tasks and applying this framework to diverse, complex tasks to further advance the development of human-level AI.

KEYWORDS

artificial general intelligence (AGI), operant conditioning, non-axiomatic reasoning system (NARS), machine psychology, adaptive learning

1 Introduction

Artificial General Intelligence (AGI) is the task of building computer systems that are able to understand or learn any intellectual task that a human being can. This type of AI is often contrasted with narrow or weak AI, which is designed to perform a narrow task (e.g., facial recognition or playing chess). There are several diverse research approaches

to AGI including brain-based approaches (e.g., Hawkins, 2021), projects that aim to implement different cognitive functions separately (e.g., Laird, 2019), and principle-based approaches (e.g., Hutter, 2004; Wang, 2013). Recently, Large Language Models like GPT-4 have also been introduced as a potential pathway towards achieving more generalizable AI systems (Bubeck et al., 2023).

One major challenge in contemporary AGI research is the lack of coherent theoretical frameworks (Wang, 2012; 2019). This scarcity of unified models to interpret and guide the development of AGI systems seems to have led to a fragmented landscape where researchers often work in isolation on narrowly defined problems. Coherent research frameworks could also provide a roadmap and evaluation criteria for AGI development, fostering more collaborative and interdisciplinary efforts. The fact that AGI research has not progressed as rapidly as some had hoped could very well be attributed to the lack of these comprehensive frameworks and the absence of standardized benchmarks for measuring progression towards AGI capabilities.

This work aims to address this challenge by proposing a novel framework that outlines key milestones and metrics for evaluating progress in the field of artificial general intelligence (AGI). A fundamental assumption in this work is that *adaptation* is at the heart of general intelligence. Adaptation is typically divided into *ontogenetic adaptation*, which involves the changes that occur with an organism over its lifespan, and *phylogenetic adaptation*, which refers to the evolutionary changes that occur across generations within a species.

Learning has within the field of learning psychology, been equated with ontogenetic adaptation, where an individual's experiences directly impact its capabilities and behaviors (De Houwer et al., 2013). *Operant conditioning* is one type of learning that involves adaptation in the form of behavioral changes due to consequences of actions. Given the enormous amount of empirical progress generated by operant conditioning research in learning psychology, the principle of operant conditioning and its associated research tradition could be a guiding principle for AGI research.

One particular approach to building AGI is the Non-Axiomatic Reasoning System (NARS) (Wang, 2013; 2022). NARS is an adaptive reasoning system that operates on the principle of insufficient knowledge and resources, a condition that is often true for real-world scenarios. Hence, NARS is a principle-based approach that aims to address the challenges of building AGI systems that can operate effectively in dynamic and unpredictable environments (Wang, 2019). There are several NARS implementations available. One implementation is OpenNARS for Applications (ONA), that is designed to provide a practical framework for integrating NARS into various applications, with a particular focus on robotics (Hammer and Lofthouse, 2020; Hammer et al., 2023). ONA is built with sensorimotor reasoning at its core, enabling it to process sensory data in real-time and respond with appropriate motor actions. Sensorimotor reasoning, as implemented in ONA, permits the system to make sense of the world much in the same way as biological organisms do, by directly interacting with its environment and learning from these interactions. The fact that NARS systems are focused on adaptation, and that ONA has a strong emphasis on sensorimotor capabilities, suggests that they are particularly well-suited for implementing the principle of operant conditioning.

This work presents Machine Psychology, an interdisciplinary framework for advancing AGI research. It integrates principles from learning psychology, with the theory and implementation of NARS. Machine Psychology starts with the assumption that adaptation is fundamental to intelligence, both biological and artificial. As it is presented here, Machine Psychology is guided by the theoretical framework of learning psychology, and the principle of operant conditioning in particular. A sensorimotor-only version of ONA (Hammer, 2022) is used to demonstrate the feasibility of using these principles to guide the development of intelligent systems. One way to describe the integration of operant conditioning and NARS presented in this paper is that the ability to learn and adapt based on feedback from the environment, is implemented using sensorimotor reasoning that is the core of ONA. An analogy is that a neurobiological explanation of operant conditioning could be argued to be part of a biological basis for adaptive behaviors observed in many species (Brembs, 2003), similarly, the implementation within ONA using temporal and procedural inference rules offers an alternative explanation of the core of adaptive behavior and cognition.

We evaluate the Machine Psychology framework by carrying out three operant learning tasks with NARS. The first is a simple discrimination task in which NARS needs to learn, based on feedback, to choose one stimulus over another, demonstrating a fundamental aspect of learning based on the consequences of actions. The second task is more complex than the first in that the conditions of the experiment are changed midway through the task, requiring NARS to adapt its choice strategy based on the new conditions. The third experiment is a conditional discrimination task, where NARS is presented with pairs of stimuli, and must learn to select the correct stimulus based on a conditional cue that changes throughout the task, requiring an increased level of adaptability. Methods from learning psychology are used to design the experiments and guide the evaluation of the results. We explain the results by describing how the sensorimotor reasoning used by ONA enables it to adaptively modify its behavior based on the consequences of its actions. The Machine Psychology framework is demonstrated to provide a coherent experimental approach to studying the core of learning and cognition with artificial agents, and also offers a scalable and flexible framework that potentially could significantly advance research in Artificial General Intelligence (AGI).

The paper is organized as follows. Section 2 presents a background on the principles of operant conditioning and its significance in learning psychology. Section 3 introduces NARS with a focus on its foundational concepts. Section 4 describes the architecture of OpenNARS for Applications with a particular focus on its sensorimotor reasoning abilities. Section 5 discusses related work to our approach. Section 6 presents the Machine Psychology framework and how it integrates operant conditioning principles with NARS. Section 7 describes the details of the methods and experimental setup used in the evaluation of our approach. Section 8 presents the results from the experiments. Section 9 concludes the paper and outlines how the Machine Psychology framework could be used to further advance the field of AGI.

2 Operant conditioning

The work presented in this paper takes a *functional* approach to learning and adaptation (De Houwer and Hughes, 2020), and to science in general. Such approach to learning is rooted in the principles of behaviorism, which emphasizes the role of environmental interactions in shaping behavior, rather than mechanistic explanations of how internal processes affect behavior. It stems particularly from the work of B. F. Skinner, who laid much of the groundwork for understanding how consequences of an action affect the likelihood of that action being repeated in the future (Skinner, 1938). Skinner was influenced by physicist and philosopher Ernst Mach, who emphasized the use of *functional relations* in science to describe relations between events, rather than using a traditional mechanistic causal framework (Chiesa, 1994).

In the functional learning psychology tradition, learning (as ontogenetic adaptation) is defined as a change in behavior due to regularities in the environment (De Houwer et al., 2013). Several types of learning can be classified under this perspective. Operant conditioning is defined as a change in behavior due to regularities between behavior and stimuli (De Houwer et al., 2013). Other types of learning can be defined based on other the regularities in operation. A few comments regarding these definitions of learning and operant conditioning follow, as clarified by De Houwer et al. (2013). First, in line with Skinner (1938) *behavior* is defined very broadly, encompassing any observable action or response from an organism. This includes responses that are only in principle observable, such as internal physiological changes, neural processes or cognitive events. In addition to this, behavior is defined to always be a function of one or more stimuli, while a response is just an observable reaction. Second, *regularities* are defined to be any patterns of events or behavior that go beyond a single occurrence. This can be the same events happening repeatedly, or two or more events or behaviors happening at the same time. Third, the definition signals a particular view of causality (“due to”). As highlighted above, this research tradition emphasizes *functional relations* between environmental regularities and changes in behavior. This implies that learning, from this perspective, cannot be directly observed, but must be inferred from the systematic changes in behavior in response to modifications in the environment (De Houwer et al., 2013). Such inferences do depend on an observer, whose scientific goals and theoretical orientations shape the interpretation. An example follows that aims to clarify this definition further.

2.1 The three-term contingency

To describe behavior as an interaction between the organism and its environment, Skinner introduced the concept of the three-term contingency, which consists of a discriminative stimulus (Sd), a response (R), and a resultant stimulus (Sr) (Skinner (1953); De Houwer and Hughes (2020)). In some contexts, the terms antecedent, behavior, and consequence are used to reflect the same triadic relationship.

In many situations, an additional element is crucial: the *establishing operation* (EO), which is an example of a *motivating operation*. The EO modifies the efficacy of a resultant stimulus as a

reinforcer. For instance, a consequence such as food might only be effective as a reinforcer under certain conditions, for example, after food deprivation - which would then be the establishing operation. Hence, the EO could be included as a fourth term in the three-term contingency description.

Furthermore, a *conditional discriminative stimulus* can modify the contingency based on additional contextual factors. This stimulus signals whether the relation between the discriminative stimulus, the response, and the resultant stimulus would hold or not, adding another layer of complexity to the model (Lashley (1938); Stewart and McElwee (2009)). The functions of the other terms would be conditional that stimulus, which explains the name conditional discriminative stimulus.

2.2 An example of operant conditioning

Imagine a rat in an experimental chamber used to study behavior. The chamber contains a small loudspeaker, a lever that can be pressed, and a water dispenser where the delivery of water is controlled by the researcher. The researcher aims to shape the rat into turning around when techno music is played and pressing the lever when classical music is played. To do this, the researcher uses an operant conditioning procedure. Before the experiment, the rat has been deprived of water for a short period, serving as the establishing operation (EO), which increases the effectiveness of water as a reinforcer.

Initially, any tendency to turn around when techno music is played (the discriminative stimulus, Sd) might be followed by the delivery of water (the resultant stimulus, Sr). This makes the behavior (R) more likely to occur in the future under similar conditions, meaning that the water functioned as a reinforcer. Conversely, when classical music (another Sd) is played, and the rat presses the lever (R), the delivery of water (Sr) follows the behavior as well. Over time, the rat learns to turn around or press the lever based on the type of music that is playing. A video of a rat performing in this experiment can be found online (WMU Rat Lab, 2009).

We could also have imagined enhancing the procedure with adding a light that could be on or off, signaling if the relation between music type, the rat's behavior, and water would hold or not. Since the function of the other terms would be conditional on the state of the light, the light would function as conditional discriminative stimulus.

This example can be considered an effective demonstration of operant conditioning. The behavior in this example is a function of both the music and the water. It illustrates the point made above that behavior studied from this perspective is not just an isolated motor action, but is also significantly influenced by the surrounding environment and the consequences that follow the behavior. The regularities in operation are reoccurring patterns of behavior and stimuli, for example, lever pressing and the delivery of water, but also a regularity regarding classical music and lever pressing. To make a causal statement about learning, we would need to observe the rat before and after interacting with these stimuli. Before the interaction, we might, for example, observe the rat exploring the cage or do random actions when different types of music was playing. After the interaction however, if we observe the behaviors described in the example, this would be a clear change in behavior

- from, for example, cage exploring when classical music is playing to lever pressing when the same music is playing. If we could argue that the change in behavior is due to the procedural arrangements (regularities), then we could potentially claim that this qualifies as an instance of learning. The type of learning it would indicate is operant conditioning since the regularities involved were between responses and stimuli (rather than, for example, a repeated pairing of stimuli as with classical conditioning). More specifically, it would be an instance of *positive reinforcement*, a kind of operant conditioning that involves an increase in target behavior due to the consequences.

2.3 Three levels of analysis

A learning situation such as the one illustrated in the example can be analyzed on three levels using learning psychology: 1) The descriptive level (or level of procedure), 2) The functional level (or the level of effect), and 3) The cognitive level (or the level of mechanism) (De Houwer and Hughes, 2020). At the descriptive level, the procedural arrangement is from the perspective of the researcher. The different sounds is used to signal if a relationship between behaving in a certain way and water holds. This is a description of procedures initiated by the researcher. It does not mean that the rat has learned based on these arrangements. The functional level however, is closer to describing the relations from the rat's perspective. If the rat turns around if and only if the techno music is playing, then that music functions as a cue for that behavior. Similarly, it is only if the delivery of water increased the behavior, that it functions as a reinforcer. If there is no change in lever pressing or turning due to the water being delivered, then that consequence has no effect. Finally, the cognitive level can be used to describe certain mental mechanisms, like association formations, that could explain how the operant learning processes take place (De Houwer and Hughes, 2020).

The importance of distinguishing these levels cannot be overstated. Just using an operant conditioning procedure (like the one above), does not mean that the subject learns in the form of operant conditioning. When doing functional learning research, we arrange procedures and study the effect on behavior change. Learning, from this perspective, is hence defined on the functional level. As stated above, a term such as reinforcement is also defined as an effect rather than a mechanism. This also means that explanations on the cognitive/mechanistic level are not part of the learning definition (De Houwer et al., 2013). It also opens up for different kinds of explanations in terms of mechanisms - for example, propositional networks as something different from association-based learning theories.

2.4 Where is the organism?

In a functional analysis of behavior, it is an interaction between organism and environment that is being analyzed. Hence, it is not the organism itself in isolation that is of interest. Technically, it is interactions between stimulus functions and response functions that are being studied, for example, an interaction between seeing a lever and pressing it, or hearing techno music and reacting to it. There are other conceptual schemas in functional learning psychology than

the three-term contingency that takes into account the complexity of these interactions (Hayes and Fryling, 2018), but for this paper, what has been presented above is a sufficient conceptual framework. Importantly though, in functional learning research, these relations between procedural arrangements and behavior change does depend on an organism in that they enable response functions. This does not in any way mean that the organism causes behavior. Rather, the organism can be considered a *participant* in the arrangements (Roche and Barnes, 1997).

Recently, De Houwer and Hughes (2022) extended their conceptual work on learning beyond that of organisms, for example, to also incorporate the study of learning with genes, groups, and machines. At the descriptive level of analysis, they replace the term response with that of *state transition*. In the example above with the rat, a change from exploring the cage to pressing the lever could be described as such a state transition (moving from a state of exploration to a state of lever pressing). Importantly, states, as defined from this perspective, are used to describe state transitions. A behavior is then defined as a state transition in relation to one or more stimuli, for example, in relation to the music being played. This is once again a functional definition - the behavior is a function of stimuli. Learning is still defined as changes in behavior (state transitions in relation to stimuli) that occurs due to regularities (De Houwer and Hughes, 2022).

Based on this conceptual change a *system* is defined as a construct from the perspective of an observer, as sets of states that can be used to describe change. In common language we often refer to a rat "being a system", but technically it is rather that the system is a collection of topographical descriptions of a rat's physiological responses. With computer systems they could typically be described as collections of interdependent systems (Hayes and Fryling, 2018). For example, a robot taking part of an experimental task, might, for example, be described as hardware movements (like the robot arm) that are dependent on sensory equipment and on software interpreting those sensory inputs.

2.5 Human-level intelligence from an operant perspective

All above examples involve learning that are animal-level in the sense that they could be observed with an animal like a rat. The fact that operant conditioning can be observed with both humans and animals does not, however, mean that these processes are irrelevant for achieving human-level intelligence with artificial systems. On the contrary, we would argue that mechanisms enabling operant conditioning at the core (as with OpenNARS for Applications; Hammer, 2022) could very well be integral to the development of complex cognitive behaviors with AGI systems.

However, there might be limitations to the purely operant account as presented by Skinner. Already in the 1950s, critiques of Skinner's theories emerged, particularly in the context of language acquisition, arguing that verbal behavior requires more than just operant conditioning (Chomsky, 1959). Given this, the development of human-level AI from the perspective of operant conditioning could risk to be limited when it comes to higher-order functions such as language. These critiques have been thoroughly addressed by contemporary functional learning

theories of language and cognition, such as Relational Frame Theory (RFT; Hayes et al., 2001).

RFT posits that the foundation of human language and cognition lies in the ability to relate objects and events in arbitrary ways, an ability referred to as Arbitrarily Applicable Relational Responding (AARR; Hayes et al., 2001; Johansson, 2019). AARR allows individuals to respond to one stimulus in terms of another based on arbitrary contextual cues, rather than solely on the physical properties of the stimuli themselves. For example, learning to relate to an object and a spoken word as equivalent, despite having no inherent physical similarity, showcases this ability. Or, responding to a small coin as more valuable than a larger one, is another demonstration of AARR.

Importantly, such patterns of AARR (called “relational frames” in RFT) are assumed to be operant behaviors in themselves, which are learned through interaction with the environment and are subject to reinforcement (Hayes et al., 2021). Learning to derive relations in accordance with, for example, similarity, opposition, comparison, etc., seems to enable the development of complex cognitive skills such as language understanding, problem-solving, and abstract reasoning. Hence, from the perspective of RFT, intelligence as a whole can be viewed as a collection of interrelated relational frames that are dynamically shaped and modified through continuous engagement with one’s environment (Cassidy et al., 2016; Hayes et al., 2021).

The implications of RFT for AGI research are profound. Since AARR is assumed to be learned behavior, it suggests that AGI systems could potentially achieve human-like intelligence through extensive training on relational tasks. There is a large amount of experimental RFT studies regarding training of relational framing abilities with humans across a variety of contexts and populations, providing a rich dataset to inform AGI development (Dixon et al., 2014; Cassidy et al., 2016). Also, an implication of an RFT perspective is the necessity for AGI systems to have mechanisms to learn relational frames from interactions with the environment in a manner similar to how humans learn throughout their lifetime. This means that AGI systems must be designed with architectures capable of not just operant conditioning in a traditional sense, but also with the ability to derive and apply relational frames dynamically.

A roadmap to AGI from the perspective of functional learning psychology and RFT would clearly emphasize operant conditioning abilities at the core, as suggested in this paper. Furthermore, it would advocate for trainings of increasingly complex AARR, in order to foster the development of advanced cognitive abilities (Johansson, 2020). Hence, such a roadmap could provide a clear specification of the requisite stages and milestones necessary for the development of AGI, aligning with the principles of Relational Frame Theory (RFT) (Hayes et al., 2001; 2021).

3 Non-axiomatic reasoning systems

A Non-Axiomatic Reasoning System (NARS) is a type of artificial intelligence system that operates under the assumption of insufficient knowledge and resources (AIKR) (Wang, 1995; 2006; 2013). The AIKR principle dictates that the system must function effectively despite having limited information and

computational resources, a scenario that closely mirrors real-world conditions and human cognitive constraints.

All NARS systems implement a Non-Axiomatic Logic (NAL) (Wang, 2013), a term logic designed to handle uncertainty using experience-grounded truth values. Most NARS systems also makes use of concept-centric memory structure, which organizes the system’s memory based on terms and subterms from the logic statements, leading to a more effective control of the inference process. Furthermore, all NARS systems use a formal language *Narsese*, that allows encoding of complex information and communication of NAL sentences within and between NARS systems.

3.1 Core principles of NARS

NARS systems are built on a few key concepts that distinguish them from traditional AI systems (Wang, 2022).

1. **Adaptation Under AIKR:** Unlike systems that assume abundant knowledge and resources, NARS thrives under constraints. It manages finite processing power and storage, operates in real-time, and handles tasks with varying content and urgency. This adaptability ensures that NARS remains relevant in dynamic and unpredictable environments.
2. **Experience-Based Learning and Reasoning:** Central to NARS is its concept-centered representation of knowledge. Concepts in NARS are data structures with unique identifiers, linked through relations such as inheritance, similarity, implication, and equivalence. These relationships are context-sensitive and derived from the system’s experiences, allowing NARS to continuously update and refine its knowledge base as it encounters new information.
3. **Non-Axiomatic Logic:** Traditional AI often relies on axiomatic systems where certain truths are taken as given. In contrast, NARS employs non-axiomatic logic, where all knowledge is subject to revision based on new experiences. This approach supports a variety of inference methods, including deduction, induction, abduction, and analogy, enabling NARS to reason in a manner that is both flexible and grounded in empirical evidence.

3.2 Problem-solving and learning

NARS processes three types of tasks: incorporating new knowledge, achieving goals, and answering questions. It uses both forward and backward reasoning to handle these tasks, dynamically allocating its limited resources based on task priorities. This approach, known as case-by-case problem-solving, means that NARS does not rely on predefined algorithms for specific problems. Instead, it adapts to the situation at hand, providing solutions that are contextually appropriate and continuously refined.

Learning in NARS is a self-organizing process (Wang, 2022). The system builds and adjusts its memory structure—a network of interconnected concepts—based on its experiences. This structure evolves over time, allowing NARS to integrate new knowledge, resolve conflicts, and improve its problem-solving capabilities.

Unlike many machine learning models that require large datasets and extensive training, NARS learns incrementally in interaction with its environment and can accept inputs at various levels of abstraction, from raw sensorimotor data to complex linguistic information.

3.3 A unified cognitive model

One of the most significant aspects of NARS is its unified approach to cognitive functions. In NARS, reasoning, learning, planning, and perception are not separate processes but different manifestations of the same underlying mechanism. This integration provides a coherent framework for understanding and developing general intelligence, making NARS a versatile tool for a wide range of AI applications (Wang, 2022).

In conclusion, NARS represents a significant departure from conventional AI paradigms by embracing the challenges of limited knowledge and resources. Its unique combination of non-axiomatic reasoning, experience-based learning, and adaptive problem-solving positions NARS as a robust model for advancing artificial general intelligence.

4 OpenNARS for applications

OpenNARS for Applications (ONA) is a highly effective implementation of a NARS, designed to be suitable for practical applications such as robotics (Hammer and Lofthouse, 2020). At the core of ONA lies sensorimotor reasoning, which integrates sensory processing with motor actions to enable goal-directed behavior under conditions of uncertainty and limited resources. ONA differs from other NARS systems in several key aspects, including.

1. **Event-Driven Control Process:** ONA incorporates an event-driven control mechanism that departs from the more probabilistic and bag-based approach used in traditional NARS implementations, such as OpenNARS (Lofthouse, 2019). This shift allows ONA to prioritize processing based on the immediacy and relevance of incoming data and tasks. The event-driven approach is particularly advantageous in dynamic environments where responses to changes must be timely and context-sensitive.
2. **Separation of Sensorimotor and Semantic Inference:** Unlike other NARS models that often blend various reasoning functions, ONA distinctly separates sensorimotor inference from semantic inference (Hammer and Lofthouse, 2020). This division allows for specialized handling of different types of reasoning tasks—sensorimotor inference can manage real-time, action-oriented processes, while semantic inference deals with abstract, knowledge-based reasoning. This separation helps to optimize processing efficiency and reduces the computational complexity involved in handling diverse reasoning tasks simultaneously.
3. **Resource Management:** ONA places a strong emphasis on managing computational resources effectively, adhering to the Assumption of Insufficient Knowledge and Resources (AIKR). It is designed to operate within strict memory and

processing constraints, employing mechanisms like priority-based forgetting and constant-time inference cycles. These features ensure that ONA can function continuously in resource-limited settings by efficiently managing its cognitive load and memory usage.

4. **Advanced Data Structures and Memory Management:** ONA utilizes a sophisticated system of data structures that include events, concepts, implications, and a priority queue system for managing these elements. This setup facilitates more refined control over memory and processing, prioritizing elements that are most relevant to the system's current goals and tasks. It also helps in maintaining the system's performance by managing the complexity and volume of information it handles.
5. **Practical Application Focus:** The architectural and control changes in ONA are driven by a focus on practical application needs, which demand reliability and adaptability. ONA is tailored to function effectively in real-world settings that require autonomous decision-making and adaptation to changing environments, making it more applicable and robust than its predecessors for tasks in complex, dynamic scenarios (Hammer and Lofthouse (2020)).

4.1 The architecture of ONA

ONA's architecture is composed of several interrelated components that work together to process sensory input, manage knowledge, make decisions, and learn from experience. These components are designed to handle the dynamic and uncertain nature of real-world environments, ensuring that the system can adapt and respond effectively (Hammer, 2022). The architecture is illustrated in Figure 1. ONA's architecture has a number of key components: 1) Event Providers, 2) FIFO Sequencer, 3) Cycling Events Queue, 4) Concept Memory, 5) Sensorimotor Inference Block, and 6) Declarative Inference Block. Each of these components plays a crucial role in ONA's operation, as described in detail below.

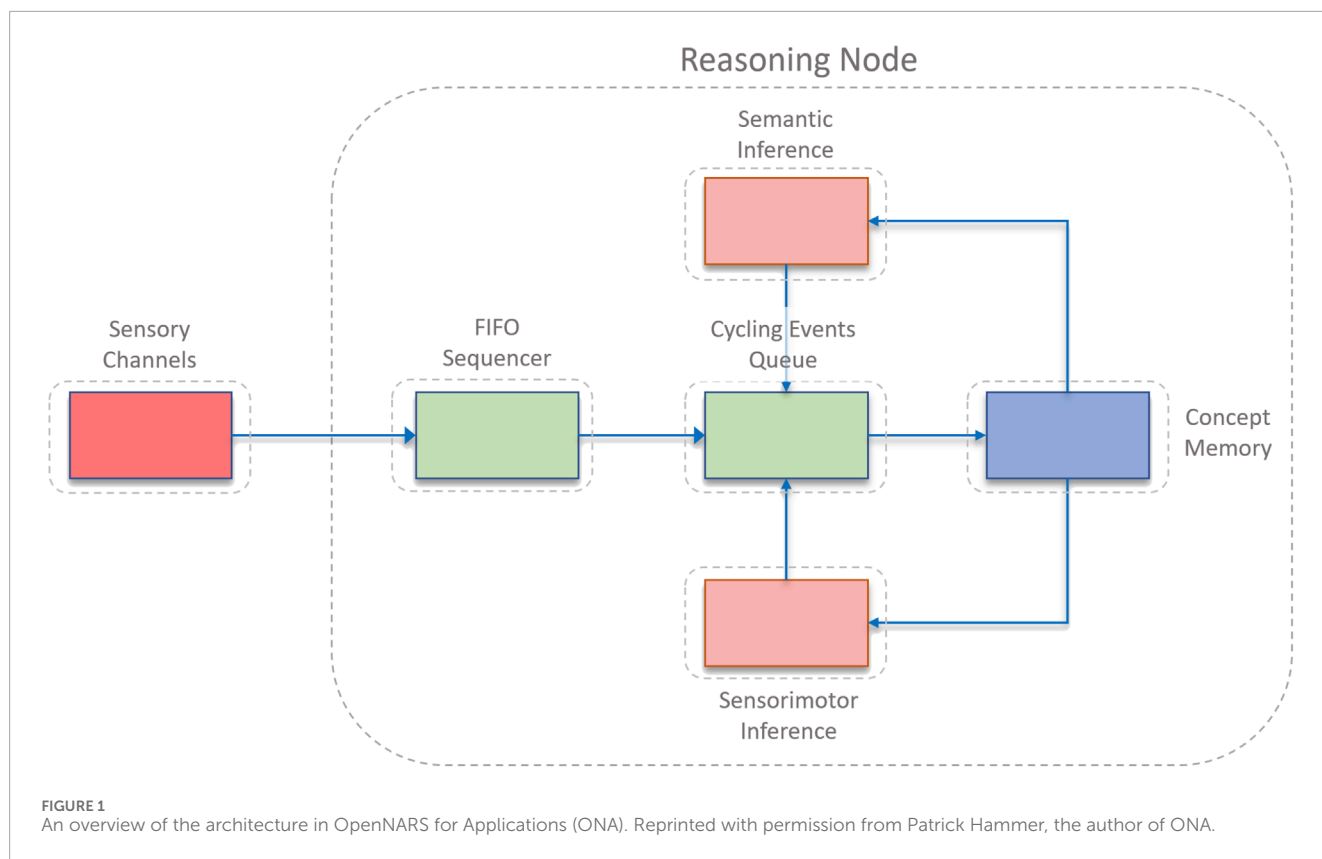
4.1.1 Event providers

Event providers are responsible for processing sensory inputs from various modalities, converting raw data into structured statements that the reasoning system can interpret. Each event provider is specialized for different types of sensory information, such as visual, auditory, or tactile data. These providers ensure that all relevant environmental information is captured and encoded as events, which are then fed into the system for further processing. The main functionality can be summarized as follows.

- **Sensor Data Processing:** Event providers preprocess raw sensor data to filter noise and extract meaningful information.
- **Event Encoding:** The processed data is encoded into statements or events that can be understood by the ONA system.

4.1.2 FIFO sequencer

The FIFO (First-In-First-Out) Sequencer maintains a sliding window of recent events. This component is essential for building and strengthening temporal implication links, which are used to



understand the sequence of events and their relationships over time. By keeping track of the recent history, the FIFO Sequencer allows ONA to form hypotheses about temporal patterns and causal relationships. As a note, in recent versions of ONA, the FIFO was removed and replaced by an explicit temporal inference block. This design is however not yet described in any scientific publications, and therefore the design with the FIFO has been described. The main functionality can be summarized as follows.

- **Event Sequencing:** It organizes events in a chronological order, maintaining a window of the most recent events.
- **Temporal Implications:** Builds and strengthens links between events based on their temporal proximity and sequence.

4.1.3 Cycling Events Queue

The Cycling Events Queue is a priority queue that serves as the central attention buffer of the system. All input and derived statements enter this queue, but only a subset can be selected for processing within a given timeframe due to the fixed capacity of the queue. This mechanism ensures that the most relevant and urgent information is processed first, while less critical information is discarded or delayed. The main functionality can be summarized as follows.

- **Priority Management:** Events are prioritized based on their importance and relevance to current goals.
- **Attention Focus:** Ensures that the system's limited processing resources are focused on the most critical tasks.

4.1.4 Concept Memory

Concept Memory acts as the long-term memory of the ONA system. It stores temporal hypotheses and supports their strengthening or weakening based on prediction success. This memory component allows ONA to retain knowledge over long periods, enabling cumulative learning and the ability to recall past experiences to inform current decision-making. The main functionality can be summarized as follows.

- **Hypothesis Management:** Stores and manages temporal and procedural hypotheses about the environment.
- **Evidence Accumulation:** Strengthens or weakens stored hypotheses based on new evidence and prediction outcomes.

4.1.5 Sensorimotor Inference Block

The Sensorimotor Inference Block is responsible for handling decision-making and subgoal processes for goal events selected from the Cycling Events Queue. This component invokes algorithms for goal achievement, generating actions or subgoals that guide the system's behavior towards fulfilling its objectives. The main functionality can be summarized as follows.

- **Decision Making:** Selects the best actions to achieve current goals based on stored knowledge and recent events.
- **Subgoaling:** Decomposes complex goals into manageable subgoals, facilitating step-by-step achievement of objectives.

4.1.6 Declarative Inference Block

The Declarative Inference Block is responsible for higher-level reasoning tasks such as feature association, prototype formation, and relational reasoning. It utilizes human-provided knowledge to enhance the system's understanding of the environment and improve its reasoning capabilities. Though not utilized in the specific experiments described in the paper, this block is crucial for applications requiring complex knowledge integration and abstract reasoning. The main functionality can be summarized as follows.

- **Feature Association:** Links related features and concepts based on observed patterns and external knowledge.
- **Relational Reasoning:** Understands and reasons about relationships between different concepts and entities.

4.2 The operations of ONA

The main operations of ONA will be described below.

4.2.1 Truth value calculation

Truth values in ONA are based on positive and negative evidence supporting or refuting a statement, respectively. The system uses two measures: frequency (the ratio of positive evidence to total evidence) and confidence (the ratio of total evidence to total evidence plus one). This approach allows ONA to represent degrees of belief, accommodating the inherent uncertainty in real-world information.

The calculation of frequency and confidence is conducted as follows. Frequency: $f = \frac{w^+}{w}$, and Confidence: $c = \frac{w}{w+1}$, where w is the total amount of evidence, and w^+ is the positive evidence.

These values are used to evaluate the truth of implications and guide decision-making processes, ensuring that actions are based on the most reliable and relevant information available.

4.2.2 Implications and learning

ONA forms temporal and procedural implications through induction and revises them based on new evidence. Temporal implications represent sequences of events, while procedural implications represent action-outcome relationships. Learning involves accumulating positive and negative evidence for these implications and adjusting their truth values accordingly. Also if an implication exists (for example, $\langle \langle A1 \rightarrow [\text{left}] \rangle \ \&/\ \wedge \text{left} \rangle \Rightarrow G$), and $A1$ was observed followed by $\wedge \text{left}$, an assumption of failure will be applied to the implication for implicit anticipation. This means, if the anticipation fails, the truth of the implication will be reduced by the addition of negative evidence, via an implicit negative G event, while the truth will increase due to positive evidence in case G happened.

The learning process at the core consists of.

- **Event Sequences:** Implications are formed when related events occur within the sliding window maintained by the FIFO Sequencer.
- **Evidence Update:** Positive and negative evidence is accumulated and used to revise the truth values of implications, ensuring they reflect the system's experiential knowledge.

4.2.3 Decision making and subgoaling

ONA's decision-making process is goal-driven, leveraging its knowledge of temporal and procedural implications to select actions or generate subgoals. The system evaluates the desire value of goals and subgoals, prioritizing them based on their likelihood of success and relevance to current objectives. The decision process can be described as follows.

- **Goal Deduction and Evaluation:** Determines the most desirable actions or subgoals based on stored implications and recent events.
- **Subgoal Generation:** Breaks down complex goals into smaller, manageable subgoals, facilitating efficient achievement through step-by-step actions.

4.2.4 Motor babbling

To trigger executions when no procedural knowledge yet exists, ONA periodically invokes random operations, a process called *Motor Babbling*. This enables ONA to execute operations despite any procedural knowledge that applies. Without this ability, ONA would not be able to do its initial steps of learning procedural knowledge (Hammer and Lofthouse, 2020).

4.3 Conclusion

The architecture of ONA integrates various components that collectively enable it to reason, learn, and make decisions under conditions of uncertainty and resource constraints. By demonstrating aspects of natural intelligence, such as continuous learning and goal-driven behavior, ONA offers a robust framework for developing intelligent systems capable of adapting to the complexities of real-world environments.

5 Related work

While we are not aware of any other attempt to integrate functional learning psychology with the Non-Axiomatic Reasoning System (NARS), there are several approaches that aim to implement the "biological basis" of operant conditioning using computational modeling or similar approaches. Importantly, it seems like most of these attempts take a mechanistic approach to operant conditioning, rather than a functional approach (as in this paper). Reinforcement learning, particularly model-free methods like Q-Learning and Deep Q-Networks (DQN), has gained significant attention for its ability to learn optimal policies through interactions with the environment (Mnih et al., 2015). These methods rely on the Markov property, where the next state depends only on the current state and action, simplifying the learning process but also limiting the system's ability to handle non-Markovian environments.

ONA diverges from RL by adopting a reasoning-based approach grounded in Non-Axiomatic Logic (NAL). Unlike RL, which optimizes a predefined reward function, ONA emphasizes real-time reasoning under uncertainty, adapting to insufficient knowledge and resources (Wang, 2013). This allows ONA to handle complex, non-Markovian environments more effectively. While RL methods struggle with sparse rewards and require extensive data to learn,

ONA leverages its reasoning capabilities to infer causal relationships and plan actions based on partial knowledge, making it more data-efficient (Hammer, 2022).

In summary, while reinforcement learning remains a powerful tool for specific, well-defined tasks, ONA offers a robust alternative for more complex, real-time applications. Its integration of reasoning under uncertainty, goal-driven learning, and adaptability positions it as a significant advancement in the quest for generalizable and resilient AI systems (Hammer, 2022).

6 Machine Psychology

Machine Psychology is an interdisciplinary framework for advancing AGI research. It aims to integrate principles from operant learning psychology (as described in Section 2), with the theory and implementation of NARS (as described in Sections 3 and 4). At the core of the integration is the assumption that adaptation is fundamental to intelligence, both biological and artificial.

Generally, Machine Psychology can be said to be a *functional* approach (as defined in Section 2 to the problem of building an AGI system). With this, we mean that the Machine Psychology framework enables the possibility to not only study functional relations between changes in the environment and changes in behavior (as in operant psychology), but also to study functional relations between mechanisms and changes in behavior. Hence, both experience of the system, and its mechanisms could in principle be manipulated.

In the case with studying operant conditioning with NARS, it means that it is indeed possible to both manipulate the system's experience, but also, in principle, to manipulate the mechanisms that are available (or not) during an experimental task.

This interdisciplinary approach might be likened to Psychobiology, that integrates psychology and biology (Dewsbury, 1991). Psychobiology is an interdisciplinary field that integrates biological and psychological perspectives to study the dynamic processes governing behavior and mental functions in whole, integrated organisms. It emphasizes the interaction between biological systems, such as the nervous and endocrine systems, and psychological phenomena, such as cognition, emotion, and behavior. This approach allows for the dual manipulation of factors related to both experience (as in psychology) and biological processes (as in biology) within a unified framework. By doing so, psychobiology provides a comprehensive understanding of how environmental and experiential factors can influence biological states and how biological conditions can shape psychological experiences, thus bridging the gap between the two domains to offer holistic insights into human and animal behavior (Dewsbury, 1991; Ritz and von Leupoldt, 2023).

Hence, one way to describe Machine Psychology, is that it is to computer science (and particularly NARS theory), as what Psychobiology is to biology.

6.1 An interaction with NARS

Within a Machine Psychology approach to NARS, it is possible to interact with NARS as one would do with an organism

in psychological research in general. This will be illustrated in this section.

An example interaction can be described as follows, where each line ends with `:|`, indicating temporal statements. First, a nonsense symbol `A1` is presented to the left. Then, `A2` is presented to the right. After that, the event `G` is established as something desirable by the system (the `!` indicates that the system desires the event). This triggers the system to execute an operation (for example, the operation `^left`). The researcher could then provide the event `G` as a consequence, leading to a derived contingency statement by the system. The entire interaction can be described as follows (where `//` represents comments):

```
<A1 --> [left]>. :|: // A1 is presented
to the left
<A2 --> [right]>. :|: // A2 is presented
to the right
G! :|: // G is established as a desired
event
^left. :|: // ^left executed by the system
G. :|: // G is provided as a consequence
<(<A1 --> [left]> &/ ^left) =/> G> //
Derived by the system
```

This example aims to provide an example of how the researcher might interact with NARS, as if it was a biological organism. The researchers presents events, and the system responds, and the researcher once again presents an event as a consequence.

As part of this study, all interactions with ONA was done via its Python interface. Specifically, experimental designs was conducted in the Python-based open source experimental software OpenSesame, that was configured to interact with ONA (Mathôt et al., 2012; Mathôt and March 2022).

6.2 Learning psychology with NARS

Given the above example, it should be clear that interactions between NARS and its environment can be analyzed using the terms provided by functional learning psychology, as provided in Section 2. The operant learning examples are described at the descriptive level (the level of procedure), and learning effects can be described at the functional level. The analog to the cognitive/mechanistic level is the operations of the NARS system, as, for example, described in Section 4.2. The three-term contingency (as described in Section 2.1) can be used to describe relations of events, operations and consequences. In the example above, the event `<A1 -> [left]>` functions as a discriminative stimulus, `^left` is a response, and `G` functions as a reinforcer. Importantly though, the `G!` (that establishes `G` as desired event) functions as an establishing operation.

7 Methods

7.1 OpenNARS for applications

The study used a version of OpenNARS for Applications (ONA) compiled with the parameter `SEMANTIC_INFERENCE_NAL_LEVEL` set to `0`, which means that only sensorimotor reasoning

were to be used. Hence, no declarative inference rules were available during the experiments.

For all three experiments, ONA was configured at starting time in the following way:

```
* babblingops=2
* motorbabbling=0.9
* setopname 1 ^left
* setopname 2 ^right
* volume=100
```

This indicates that ONA was set to have two operators $\hat{\text{left}}$ and $\hat{\text{right}}$, and an initial chance of 90% for motor babbling.

7.2 Encoding of experimental setup

All experimental tasks were presented as temporal Narsese statements, as indicated by the $:|:$ markers below. An arbitrary goal event $G! :|:$ was presented at the end to trigger the execution of one of the two procedural operations $\hat{\text{left}}$ and $\hat{\text{right}}$ (through motor babbling or a decision). During training, feedback was given in the form of $G. :|:$ (meaning to reinforce a correct choice) or $G. :|:\{0.0\ 0.9\}$ (to indicate that the system had conducted an incorrect choice). Between each trial, 100 time steps was entered, by feeding 100 to ONA.

```
<A1 --> [sample]>. :|:
<B1 --> [left]>. :|:
<B2 --> [right]>. :|:
G! :|:
```

The first three lines are so-called inheritance statements, with properties on the right-hand side, indicating that the events $A1$, $B1$ and $B2$ are either on the left, right or at the position of a sample.

7.3 Experimental designs

In this section, the experimental designs will be detailed of the three tasks: 1) The simple discrimination task, 2) The changing contingencies task, and 3) The conditional discriminations task. The tasks are further illustrated with a few examples in [Figure 2](#).

The first experiment investigated if NARS could learn in the form of operant conditioning, specifically in the form of simple discriminations. In the experiment, three phases were used: Baseline assessment, Training (with feedback), and Testing (without feedback). In all phases, training and testing were done in blocks of trials. One trial could, for example, be that $A1$ was to the left, and $A2$ was to the right. A block contained twelve trials, with the two possible trials possible (depending on the location of $A1$ and $A2$), each presented six times in random order.

1. **Baseline:** During the baseline assessment, which was three blocks, no feedback was given. This phase was included to establish a baseline probability of responding correct. It was expected that the system would respond correctly by chance in 50% of the trials.
2. **Training:** Then, the system was trained on a set of three blocks. Feedback was given when the system was correct (for example, executing $\hat{\text{left}}$ when $A1$ was to the left), and when not correct.

3. **Testing:** The system was then tested (without feedback) on three blocks, with the contingencies that previously had been trained.

The second experiment investigated if ONA could adapt to changing conditions midway through the task. Five phases were used: Baseline, Training 1 (with feedback), Testing 1 (without feedback), Training 2 (with feedback), and Testing 2 (without feedback). All blocks contained twelve trials.

1. **Baseline:** Two blocks, where no feedback was given.
2. **Training 1:** Four blocks of, where feedback was given. This phase aimed to train the system in executing $\hat{\text{left}}$ when $A1$ was to the left, and $\hat{\text{right}}$ when $A1$ was to the right.
3. **Testing 1:** Then, the system was tested over two blocks (without feedback) on what was trained the previous phase.
4. **Training 2:** This phase of four blocks aimed to train in reversed contingencies compared to the first training. That is, the phase aimed to train ONA into executing $\hat{\text{left}}$ when $A2$ was to the left (and hence $A1$ to the right), and $\hat{\text{right}}$ when $A2$ was to the right.
5. **Testing 2:** Over two blocks, the system was tested, without feedback, on the contingencies trained in the previous phase.

Finally, in the third experiment, that investigated if the system could learn conditional discriminations, three phases were used: Baseline, Training, and Testing.

1. **Baseline:** Three blocks of 12 trials, where no feedback was given.
2. **Training:** Six blocks, where feedback was given. For example, when $A1$ was the sample, and $B1$ to the left, the system was reinforced for executing $\hat{\text{left}}$.
3. **Testing:** The system was then tested, without feedback, on three blocks of 12 trials, with the contingencies that previously had been trained.

8 Results

8.1 Simple discrimination task

During baseline, the amount of correct trials ranged between 0% and 50% during the three blocks, indicating that no learning happened. In the training phase, NARS was 100% correct on all trials already in the second out of three blocks, indicating a rapid learning. Finally, in the testing, where no feedback was provided, NARS performed consistently 100% correct across all three blocks of trials. The results are illustrated in [Figure 3](#).

The average confidence values for the two target hypotheses went from 0.56 to 0.82. These two hypotheses were

```
<(<A1 --> [left]> &/ ^left) =/> G>
and
```

```
<(<A1 --> [right]> &/ ^right) =/> G>
```

The increase in average confidence value is also illustrated in [Figure 3](#).

In summary, the results indicate that ONA indeed can learn in the form of operant conditioning.

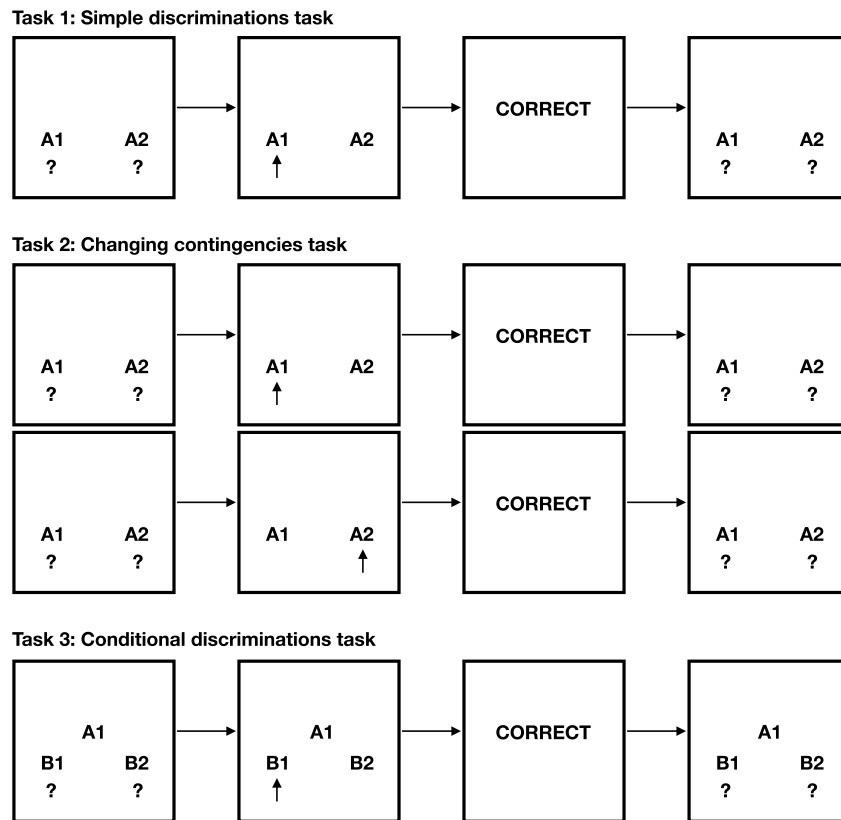


FIGURE 2 Examples from the three experimental tasks investigated.

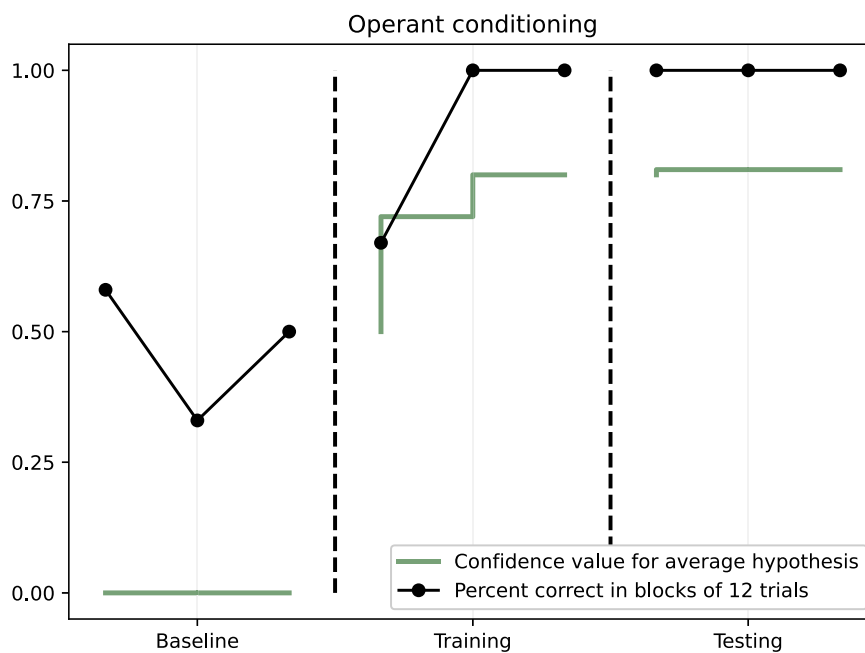
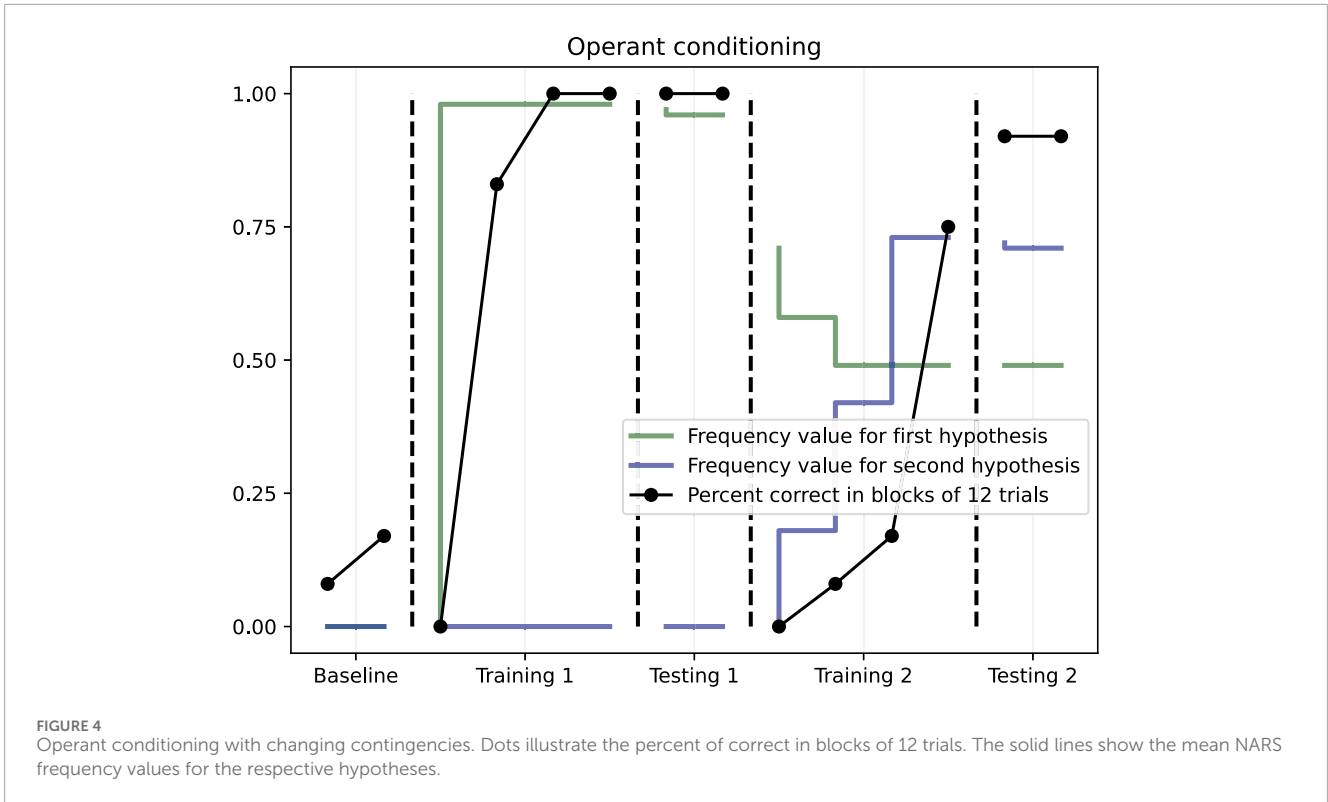


FIGURE 3 Operant conditioning. Dots illustrate the percent of correct in blocks of 12 trials. The solid line shows the mean NARS confidence value for hypotheses.



8.1.1 NARS examples from the training phase

A few example trials from the training session follows. Let's say that the system was exposed to the following NARS statements:

```
<A2 --> [left]>. :|:
<A1 --> [right]>. :|:
G! :|:
```

If it is early in the training, NARS might use Motor Babbling to execute the \hat{right} operation. Since this is considered correct in the experiment, the feedback $G. :|:$ would be given to NARS, followed by 100 time steps. Only from this single interaction, NARS would form a hypothesis using *Temporal Induction*:

```
<(<A1 --> [right]> &/ ^right) =/> G>
```

When the same situation happens again later during the training phase, ONA will not rely on motor babbling, but instead use its decision making algorithm and *Goal Deduction*, as detailed by (Hammer, 2022).

8.2 Changing contingencies task

As expected, no learning happened during the baseline phase, where NARS was less than 25% correct in both phases. In the first training phase, NARS was 100% correct after two completed blocks of 12 trials. During testing, the system was 100% correct without any feedback being present. In the second training phase, where the contingencies were reversed, the system could adapt to the change as indicated by the increase in number of correct responses over time, with 75% correct in the final block of the phase. Finally, in the second testing phase, the system's performance was 91.7% correct, indicating that a successful retraining had been conducted. The results are further illustrated in Figure 4.

To further illustrate how the NARS system was able to adapt to changing contingencies, the change in average frequency value of the two target hypotheses can be described over time. This is also illustrated in Figure 4. As seen in Figure 4, the average frequency value for the first hypothesis was close to 1.0 during the first training and testing, meaning that the system had not received any negative evidence. However, when the contingencies were reversed in the second training phase, the frequency value of the first hypothesis immediately decreased, taking the negative evidence into account. The frequency value of the second hypothesis however, did not rise above zero until the start of the second training, where the hypothesis got positive evidence for the first time.

These results do all together indicate that a NARS system can adapt in realtime in the form that is necessary when contingencies are reversed midway through a task.

8.2.1 Examples from changed contingencies

The experiment starts out similar as to the example in Section 8.1.1 during the first training phase. However, after the contingencies change, and reinforcement is not provided for executing \hat{left} and \hat{right} when A1 is to the left and right, respectively, the system is forced to readapt. For example, if the following situation is shown to the system:

```
<A1 --> [left]>. :|:
<A2 --> [right]>. :|:
G! :|:
```

The system will execute \hat{left} based on its previous learning. However, instead of $G. :|:$ as a consequence, $G. :|: \{0.0\ 0.9\}$ will be provided. An explanation of how *Revision* is used will be provided.

Before the negative feedback, the following hypothesis will have a frequency close to 1.0:

```
<(<A1 --> [left]> &/ ^left) =/> G>.
{0.98, 0.41}
```

In the above, 0.98, 0.41 means frequency = 0.98, and confidence = 0.41.

However, with the negative feedback shown above, the following hypothesis will be derived:

```
<(<A1 --> [left]> &/ ^left) =/> G>.
{0.00, 0.19}
```

Together, these two hypothesis with different truth values will be revised as follows:

```
<(<A1 --> [left]> &/ ^left) =/> G>.
{0.74, 0.48}
```

When NARS combines the positive and negative evidence, the frequency value goes down from 0.98 to 0.74, and the confidence value goes up from 0.41 to 0.48, as the system has gained even more evidence and is more confident in its conclusions.

With repeated examples similar to the above, the system will eventually go back to motor babbling, and ^right will be executed, leading to a reinforcing consequence. That will lead to the following hypothesis being formed:

```
<(<A2 --> [right]> &/ ^right) =/> G>.
```

In summary, the mechanism of *Revision*, in combination to what have been covered previously, enables the system to adapt to changing contingencies.

8.3 Conditional discriminations task

As with the previous experiments, no learning happened during the three-block baseline. During training, NARS was more than 75% correct after two completed blocks of 12 trials. In the testing, NARS performed 100% correct, without feedback, across three blocks of trials. These results are illustrated in [Figure 5](#).

The four target hypotheses were the following:

```
<((<A1 --> [sample]> &/ <B1 --> [left]>)
&/ ^left) =/> G>
<((<A1 --> [sample]> &/ <B1 --> [right]>)
&/ ^right) =/> G>
<((<A2 --> [sample]> &/ <B2 --> [left]>)
&/ ^left) =/> G>
<((<A2 --> [sample]> &/ <B2 --> [right]>)
&/ ^right) =/> G>
```

The average confidence value for these hypotheses increased from 0.13 to 0.70 during the training phase, as also illustrated in [Figure 5](#).

8.3.1 NARS examples from conditional discrimination training

A few example trials from the training session follows. Let's say that the system was exposed to the following NARS statements:

```
<A1 --> [sample]>. :|:
<B2 --> [left]>. :|:
<B1 --> [right]>. :|:
G! :|:
```

If it is early in the training, NARS might use motor babbling to execute the ^right operation. Since this is considered correct in

the experiment, the feedback G. :|: would be given to NARS, followed by 100 time steps. From this single interaction, NARS would form a hypothesis:

```
<((<A1 --> [sample]> &/ <B1 --> [right]>)
&/ ^right) =/> G>.
// frequency: 1.00, confidence: 0.15
```

Importantly, after this single trial, NARS would also form simpler hypothesis such as:

```
<(<B1 --> [right]> &/ ^right) =/> G>.
// frequency: 1.00, confidence: 0.21
<(<A1 --> [sample]> &/ ^right) =/> G>.
// frequency: 1.00, confidence: 0.16
```

This means, that if the same trial was to be presented again (all four possible trials will be presented three times in a block of twelve trials), NARS would respond ^right again, but the decision being based on the simpler hypothesis, since that hypothesis has the highest confidence value.

Let's say, that within the same block of 12 trials, the next trial to be presented to NARS was the following:

```
<A1 --> [sample]>. :|:
<B1 --> [left]>. :|:
<B2 --> [right]>. :|:
G! :|:
```

NARS would initially respond ^right, with the decision being made from the simple hypothesis <(<A1 --> [sample]> &/ ^right) =/> G>.

This would be considered wrong in the experiment, and the feedback G. :|: {0.0 0.9} would be given to NARS. This would lead to negative evidence for the simple hypothesis. If the same trial was presented again, NARS would then likely resort to motor babbling that could execute the ^left operation. Over repeated trials with feedback, the simpler hypotheses would get more negative evidence, and the confidence values of the more complex target hypotheses would increase.

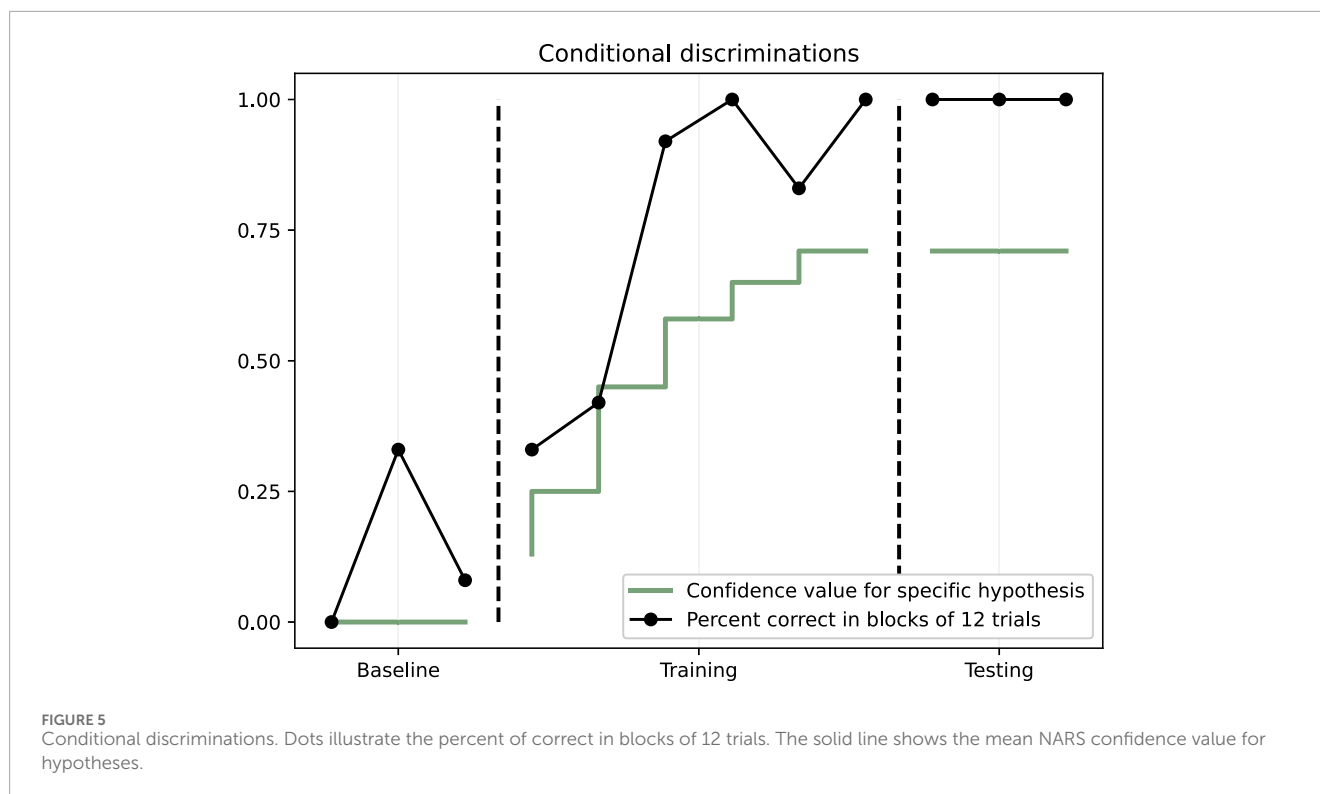
In summary, NARS can learn increasingly complex hypotheses, with repeated examples.

8.4 NARS mechanisms

Given the examples above, we will now provide further clarifications of the results in terms of mechanisms and inference rules that are implemented in ONA.

In all three tasks, the confidence increase followed from repeated examples which provide evidence to the respective target hypotheses. For this to happen and to derive the truth values, the following mechanisms in NARS were necessary.

- Temporal induction:** Given events that A1 is to the left, the ^left operation, and G, then derive positive evidence for a relation like <(<A1 --> [left]> &/ ^left) =/> G>
- Goal deduction:** Given for example, <(<A1 --> [left]> &/ ^left) =/> G> and a precondition that A1 is to the left, and the event G!, then by deduction derive that the ^left operation is to be executed.
- Motor babbling:** The ability to execute operations functions as the means for exploration in the sense that it enables the system to try out new things.



4. **Anticipation:** To derive negative evidence to a hypothesis, based on that the antecedent happened but the consequent did not. For example, $\langle \langle A1 \rightarrow [sample] \rangle \&/ \wedge \text{right} \rangle = / \rangle G \rangle$ can receive negative evidence based on anticipation.
5. **Revision:** To summarize the positive evidence and the negative evidence for a statement.

9 Discussion and conclusion

The results of this study demonstrate the feasibility and effectiveness of integrating principles from operant conditioning with the Non-Axiomatic Reasoning System (NARS) to advance the field of Artificial General Intelligence (AGI). This interdisciplinary framework, referred to as Machine Psychology, offers a novel approach to understanding and developing intelligent systems by emphasizing adaptation, a core aspect of both biological and artificial intelligence.

9.1 Summary of findings

The experiments conducted in this study aimed to evaluate the ability of NARS, specifically the OpenNARS for Applications (ONA) implementation, to perform operant conditioning tasks. The three tasks—simple discrimination, changing contingencies, and conditional discriminations—provided a comprehensive assessment of the system's learning and adaptation capabilities.

In the simple discrimination task, NARS demonstrated rapid learning, achieving 100% correct responses during the training phase and maintaining this performance in the testing phase

without feedback. This indicates that NARS can effectively learn and adapt based on positive reinforcement, a key aspect of operant conditioning.

The changing contingencies task further highlighted the system's adaptability. When the contingencies were reversed midway through the task, NARS was able to adjust its behavior accordingly, showing a significant decrease in errors and an increase in correct responses during the retraining phase. This flexibility is crucial for AGI systems operating in dynamic environments where conditions can change unpredictably.

The conditional discriminations task showcased NARS's ability to handle more complex learning scenarios. Despite the increased difficulty, the system achieved high accuracy, indicating that it can form and utilize more intricate hypotheses based on conditional cues. This capability is essential for developing AGI systems that require sophisticated cognitive skills.

9.2 Implications for AGI research

The success of NARS in these operant conditioning tasks has several important implications for AGI research. First, it validates the use of learning psychology principles, particularly operant conditioning, as a guiding framework for developing intelligent systems. The results suggest that mechanisms enabling operant conditioning are integral to the development of adaptive behaviors and cognition in AGI systems.

Second, the experiments carried out as part of this study, can be said to constitute key milestones of AGI research, as has been suggested by us elsewhere (Johansson, 2020). Operant psychology research provides examples of increasingly complex tasks, that can

be used to test the abilities of an AGI system. The use of functional learning psychology principles to guide AGI research also enable metrics to be used to evaluate AGI systems, as demonstrated in this paper.

Third, the study highlights the potential of NARS as a robust model for AGI. Unlike traditional AI systems that rely on predefined algorithms and large datasets, NARS operates effectively under conditions of insufficient knowledge and resources. This adaptability makes it well-suited for real-world applications where information is often incomplete and environments are constantly changing.

Fourth, the integration of sensorimotor reasoning with operant conditioning principles in ONA provides a scalable and flexible framework for AGI development. By demonstrating aspects of natural intelligence, such as continuous learning and goal-driven behavior, ONA offers a practical approach to building intelligent systems that can interact with and learn from their environments in real-time.

9.3 Future directions

The findings of this study open several avenues for future research. One potential direction is to explore the integration of additional cognitive and behavioral principles from functional learning psychology into NARS. Future research can be guided by operant theories of cognition, such as Relational Frame Theory, as suggested by Johansson (2019).

Another important direction is to apply the Machine Psychology framework to more complex and diverse tasks beyond the idealized examples provided in this paper. By testing NARS in various real-world scenarios, such as autonomous robotics, natural language processing, and human-computer interaction, researchers can evaluate the system's generalizability and robustness across different domains.

Additionally, further refinement of the sensorimotor inference and declarative inference components in ONA could lead to improvements in the system's performance. Enhancing the efficiency of resource management, memory structures, and event-driven control processes will be critical for scaling up the system to handle more sophisticated tasks and larger datasets.

9.4 Conclusion

In conclusion, this study demonstrates that integrating operant conditioning principles with NARS offers a promising pathway for advancing AGI research. The Machine Psychology framework provides a coherent and experimentally grounded approach to studying and developing intelligent systems. By emphasizing adaptation and learning from environmental interactions, this interdisciplinary approach has the potential to significantly advance

the field of AGI and bring us closer to achieving human-level artificial intelligence.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

RJ: Conceptualization, Formal Analysis, Investigation, Methodology, Writing—original draft, Writing—review and editing, Funding acquisition, Project administration, Visualization.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was in part financially supported by Digital Futures through grant agreement KTH-RPROJ-0146472.

Acknowledgments

The author would like to thank Patrick Hammer and Tony Lofthouse for many valuable discussions regarding the work presented in this paper.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of *Frontiers*, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Brembs, B. (2003). Operant conditioning in invertebrates. *Curr. Opin. Neurobiol.* 13, 710–717. doi:10.1016/j.conb.2003.10.002
- Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., et al. (2023). *Sparks of artificial general intelligence: early experiments with gpt-4*. *arXiv preprint arXiv:2303.12712*.
- Cassidy, S., Roche, B., Colbert, D., Stewart, I., and Grey, I. M. (2016). A relational frame skills training intervention to increase general intelligence and scholastic aptitude. *Learn. Individ. Differ.* 47, 222–235. doi:10.1016/j.lindif.2016.03.001
- Chiesa, M. (1994). *Radical behaviorism: the philosophy and the science*. (Authors Cooperative).

- Chomsky, N. (1959). Verbal behavior. *Language* 35, 26–58. doi:10.2307/411334
- De Houwer, J., Barnes-Holmes, D., and Moors, A. (2013). What is learning? on the nature and merits of a functional definition of learning. *Psychonomic Bull. & Rev.* 20, 631–642. doi:10.3758/s13423-013-0386-3
- De Houwer, J., and Hughes, S. (2020). *The psychology of learning: an introduction from a functional-cognitive perspective*. MIT Press.
- De Houwer, J., and Hughes, S. J. (2022). Learning in individual organisms, genes, machines, and groups: a new way of defining and relating learning in different systems. *Perspect. Psychol. Sci.* 18, 649–663. doi:10.1177/17456916221114886
- Dewsbury, D. A. (1991). Psychobiology. *Am. Psychol.* 46, 198–205. doi:10.1037//0003-066x.46.3.198
- Dixon, M. R., Whiting, S. W., Rowsey, K., and Belisly, J. (2014). Assessing the relationship between intelligence and the peak relational training system. *Res. Autism Spectr. Disord.* 8, 1208–1213. doi:10.1016/j.rasd.2014.05.005
- Hammer, P. (2022). “Reasoning-learning systems based on non-axiomatic reasoning system theory,” in *International workshop on self-supervised learning* (PMLR) 192, 89–107.
- Hammer, P., Isaev, P., Lofthouse, T., and Johansson, R. (2023). “Ona for autonomous ros-based robots,” in *Artificial general intelligence: 15th international conference, AGI 2022, Seattle, WA, USA, august 19–22, 2022, proceedings* (Springer), 231–242.
- Hammer, P., and Lofthouse, T. (2020). “opennars for applications’: architecture and control,” in *International conference on artificial general intelligence* (Springer), 193–204.
- Hawkins, J. (2021). *A thousand brains: a new theory of intelligence*. (Hachette UK).
- Hayes, L. J., and Fryling, M. J. (2018). Psychological events as integrated fields. *Psychol. Rec.* 68, 273–277. doi:10.1007/s40732-018-0274-3
- Hayes, S. C., Barnes-Holmes, D., and Roche, B. (2001). *Relational frame theory: a post-Skinnerian account of human language and cognition*. New York: Kluwer Academic/Plenum Publishers.
- Hayes, S. C., Law, S., Assemi, K., Falletta-Cowden, N., Shamblin, M., Burleigh, K., et al. (2021). Relating is an operant: a fly over of 35 years of rft research. *Perspect. em Análise do Comportamento* 12, 005–032. doi:10.18761/pac.2021.v12.rft.02
- Hutter, M. (2004). *Universal artificial intelligence: sequential decisions based on algorithmic probability*. Springer Science & Business Media.
- Johansson, R. (2019). “Arbitrarily applicable relational responding,” in *International conference on artificial general intelligence* (Springer), 101–110.
- Johansson, R. (2020). “Scientific progress in agi from the perspective of contemporary behavioral psychology,” in *OpenNARS workshop, artificial general intelligence (AGI-2020)*.
- Laird, J. E. (2019). *The Soar cognitive architecture*. MIT press.
- Lashley, K. S. (1938). Conditional reactions in the rat. *J. Psychol.* 6, 311–324. doi:10.1080/00223980.1938.9917609
- Lofthouse, T. (2019). “Alann: an event driven control mechanism for a non-axiomatic reasoning system (nars),” in *NARS2019 workshop at AGI*.
- Mathôt, S., and March, J. (2022). Conducting linguistic experiments online with opensesame and osweb. *Lang. Learn.* 72, 1017–1048. doi:10.1111/lang.12509
- Mathôt, S., Schreij, D., and Theeuwes, J. (2012). Opensesame: an open-source, graphical experiment builder for the social sciences. *Behav. Res. methods* 44, 314–324. doi:10.3758/s13428-011-0168-7
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., et al. (2015). Human-level control through deep reinforcement learning. *nature* 518, 529–533. doi:10.1038/nature14236
- Ritz, T., and von Leupoldt, A. (2023). Introduction to the 2022 special issue on neuroscience and psychobiology of respiration in biological psychology. *Biol. Psychol.* 176, 108478. doi:10.1016/j.biopsycho.2022.108478
- Roche, B., and Barnes, D. (1997). The behavior of organisms? *Psychol. Rec.* 47, 597–618. doi:10.1007/bf03395248
- Skinner, B. F. (1938). *The behavior of organisms: an experimental analysis*. New York, NY: Appleton-Century-Crofts.
- Skinner, B. F. (1953). *Science and human behavior*. 92904. Simon & Schuster.
- Stewart, I., and McElwee, J. (2009). Relational responding and conditional discrimination procedures: an apparent inconsistency and clarification. *Behav. Analyst* 32, 309–317. doi:10.1007/bf03392194
- Wang, P. (1995). *Non-axiomatic reasoning system: exploring the essence of intelligence*. Indiana University.
- Wang, P. (2006). *Rigid flexibility: the logic of intelligence*, 34. Springer Science & Business Media.
- Wang, P. (2012). “Theories of artificial intelligence—meta-theoretical considerations,” in *Theoretical foundations of artificial general intelligence* (Springer), 305–323.
- Wang, P. (2013). *Non-axiomatic logic: a model of intelligent reasoning* (World Scientific).
- Wang, P. (2019). On defining artificial intelligence. *J. Artif. General Intell.* 10, 1–37. doi:10.2478/jagi-2019-0002
- Wang, P. (2022). “A Unified Model of Reasoning and Learning,” in *Proceedings of the Second International Workshop on Self-Supervised Learning* (PMLR) 159, 28–48.
- WMU Rat Lab(2009). *More stimulus discrimination*. Available at: <https://youtu.be/S2SqKRCa1a8>(Accessed October 3, 2014).