THESIS

UNDERSTANDING BEHAVIORAL MECHANISMS OF PAVLOVIAN BIASES THROUGH TASK REVERSAL PARADIGMS

Submitted by

Sanjiti Sharma

Department of Psychology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2025

Master's Committee:

Advisor: Carol Seger

Michael Thomas
Ana Clara Bobadilla

ABSTRACT

UNDERSTANDING BEHAVIORAL MECHANISMS OF PAVLOVIAN BIASES THROUGH TASK REVERSAL PARADIGM

Decision-making processes are often influenced by Pavlovian biases as they subtly guide our actions and reactions unconsciously (Raab & Hartley, 2020). These learning biases are acquired overtime through experiences and result in previously neutral stimuli becoming positively or negatively valanced. Pavlovian biases can elicit automatic yet emotional responses in a variety of contexts ranging from mundane life events to life-threatening scenarios (Chen et al., 2022). Often cognitive disorders associated with corrupted decision-making can stem from unchecked Pavlovian biases. Some of these disorders include Substance Use Disorders (SUDs), depression, addictions, chronic pain as well as eating disorders (Nees et al., 2015). Importantly, Pavlovian biases often work in collaboration with instrumental learning. Instrumental learning is defined as a form of learning where actions are modified based on the consequences that they produce (Balleine & Dickinson, 1998).

This study consists of Experiment 1 (N = 90 healthy adults) as well as Experiment 2 (N = 90 healthy adults), designed to explore the interaction between Pavlovian biases and goal-directed instrumental learning by examining how each adapts to task reversal. Both experiments in this study have three key conditions: baseline, reversal of goal-directed instrumental learning and reversal of Pavlovian bias. We hypothesized that Pavlovian biases would be slow to adjust after reversal due to their reliance on inflexible learning mechanisms, whereas the more flexible, goal-directed instrumental learning system would adapt more quickly. In line with our hypothesis, we found that goal-directed instrumental learning adapted rapidly whereas Pavlovian biases initially presented hurdles in participants performance levels however, eventually demonstrated some flexibility which led participants to overcome them. This study presents evidence for flexibility demonstrated by Pavlovian biases in deterministic reversal paradigms and further propels the discourse on how these biases interact with goal-directed instrumental learning.

TABLE OF CONTENTS

INTRODUCTION

1.0 Characteristics of Learning Systems and Decision Making

Foundational research supports the presence of multiple memory and learning systems in the brain (Ashby et al., 2005; Nadel, 1992; Sherry et al., 1987). The two main divisions of memory are declarative memory systems and non-declarative memory systems which are also known as procedural memory (Squire, 2004). As the name suggests, 'declarative' memory mechanisms rely on conscious recall of events and facts which is then further divided into episodic (personal memories) and semantic memories (general knowledge and facts) (Tulving, 1985). Declarative memory is applied in various contexts that require high cognitive attention and recollection. This memory system is supported neurally through hippocampus and frontoparietal systems that assist us in applying memories to support goal-directed behaviors and make conscious decisions (Eichenbaum et al., 2000). Additionally, cerebral cortex and thalamus (specifically, the anterior and dorsomedial nuclei) also assist us in recalling explicit information (Squire, 2015). Declarative memory systems are flexible and context-sensitive allowing behavior to be modified based on the environment the agent is placed in (Eichenbaum et al., 2000).

Non-declarative memory systems have been characterized as unconscious and incremental as supported by previous research, where it was discovered that habitual learning is inflexible and not context-sensitive, highlighting its ability to persevere regardless of the environmental context (Knowlton et al., 1996). Basal ganglia and cerebellum are the two major areas that facilitate non-declarative learning processes (Foerde & Shohamy, 2011). The behavioral and neural aspects of non-declarative memory indicates two important characteristics of this form of memory: an inability to be controlled through goal-directed executive functioning and rigidity, or inflexibility (Knowlton et al., 1996). One of the key factors distinguishing declarative from nondeclarative memory is the degree of rigidity or flexibility of the memory systems (Reber et al., 1996). Rigidity in this context refers to the inability to modify behavior after the introduction of new outcomes associated with respective actions. Conversely, flexibility refers to the ability to adapt behavior to new outcomes and altered contexts. Additionally, automaticity is another one of the essential characteristics of habitual or nondeclarative learning systems. Because of automaticity, there is less cognitive load and more mental capacity available for other tasks (Graybiel, 2008). For example, once someone gets familiar with the route they need to take after work to go home, they might do so with little to no mental effort. Automaticity has evolved evolutionarily to achieve the least cognitive burden, allowing us the ability to pursue and achieve other goals that potentially need our attention (Toner et al., 2015). It has been proposed that one of these largely automatic and unconscious learning process is Pavlovian conditioning or Pavlovian biases, which allow us to rely on well-learned reactions that are crucial for our survival.

Theories of learning in decision-making have been refined over the years as well. Research suggests that, overarchingly, learning to make optimal decisions involves two primary mechanisms: automatic learning (includes mechanisms such as model-free learning, habits and classical conditioning systems) and flexible learning (encompasses model-based learning, and goal-directed behavior). Over the years, researchers have designated these two governing processes under a variety of labels such as dual-process theories (Stanovich et al., 2000), reinforcement learning framework (Dayan et al., 2013), and lastly, Pavlovian and instrumental interaction (Guitart-Masip et al., 2012). Labeled as system 1 (automatic) and system 2 (flexible), researchers Stanovich and West report that these two systems often compete and conflict to control behavior in a particular situation (Stanovich & West, 2000; Kahneman, 2011). They suggest that when these two systems are in conflict for a decision, there are no noticeable individual cognitive differences (meaning there may be homogeneity within participants responses) however, if these two systems have a variance of opinions, then we notice higher individual differences within cognitive abilities (for e.g. people with higher analytical intelligence are able to override their fast yet erroneous decisions). The longstanding conversation of these two systems has recently been revised to explore the rigidity 'system 1' exhibits on our decision-making processes. In 2019, Fabio and colleagues, challenged the idea of a dichotomous perspective on the aforementioned two systems. This research found that the use of contextual features can elicit top-down control over automatic processes, evoking flexibility (Fabio et al., 2019). Additionally, research that assessed the strength and complexities of automatic processing in cognitive control challenged the notion of exclusive bottom-up processing within these mechanisms (Kiefer, 2012). Understanding these two systems extensively through the mechanisms they employ for memory and decision-making will ensure informed judgments leading to better applications.

2.0 Classical conditioning and Pavlovian bias

Classical conditioning is a learning mechanism where an organism places value on a neutral cue due to repeated pairings with a specific outcome, which can be either aversive or appetitive (Rehman et al., 2017). This type of learning encompasses four key terms for the stimuli present during reinforcement learning, namely: unconditioned stimuli (US), unconditioned response (UR), conditioned stimuli (CS) and lastly, conditioned response (CR). A common example we can use to elaborate these terms could be the universal fear of a dentist. The pain one might feel during a dental procedure can be considered as the unconditioned stimuli (US). This pain is naturally occurring and can be considered as one's automatic response to the procedure. The fear and anxiety that accompanies that pain, can be considered as unconditioned response (UR) – natural response to the pain. When an individual is required to go into the dental office again, they might hear the dental drill (conditioned stimuli, CS) and respond with anxiety and fear – which can be classified as their conditioned response (CS). This widely known discourse on classical conditioning was first started by a Russian physiologist, Ivan Pavlov.

2.1 Pavlovian Biases

Around the 1890s, Ivan Pavlov conducted research that discovered Pavlovian biases. Publishing his first paper on classical conditioning mechanisms with animal models in 1903, Pavlov is famously recognized for his experiments on the 'dog and the bell' (Pavlov, 1903). In his experiment, he found that the dogs would start salivating (conditioned response) after hearing the bell (conditioned stimulus) as it was consistently being paired with food (unconditioned stimulus). This phenomenon then gave birth to what we now know as Pavlovian biases. Research defines Pavlovian biases as cognitive or decision-making biases that are influenced by associative reward learning (Raab & Hartley, 2019). These biases attach valence to previously neutral cues and stimuli, which then either impedes or facilitates an individual's ability to attain rewards or avoid undesirable results. Acquired through classical conditioning mechanisms, these biases can modify behavior even after the initial learning phase (Rescorla, 1988).

Pavlovian biases are said to influence decision making in two different ways: interacting with avoidance behaviors or with approach behaviors (Guitart-Masip et al., 2014). When one encounters a stimulus that was previously associated with positive outcomes such as reward, they might exhibit approach or appetitive behavior (O'Doherty, 2004). For example, you might enter a bakery due to the smell of freshly baked cookies because you have learned that the smell of cookies is associated with the enjoyment of eating them. Contrarily, encountering a cue previously attached to a negative outcome leads to avoidance or aversive behavior (Lovibond et al., 2002). For example, encountering a mountain lion on your hike might evoke an aversive or avoidance reaction assisting you to step back. As much as these rigid biases can assist us in unfamiliar situations by producing cost-effective and fast responses, they can also impede our decision-making processes (Raab & Hartley, 2020).

2.2 Acquisition & Extinction of Pavlovian biases

Pavlovian biases work through two crucial processes: acquisition and extinction (Delamater et al., 2004; Stussi et al., 2021). The acquisition phase of Pavlovian biases is via classical conditioning in which there is a repeated association of a neutral stimuli with an outcome that is either rewarding or punishing. When an individual is in the acquisition phase, they start to associate conditioned stimuli to unconditioned stimuli (food and bell respectively), evoking a conditioned response (salivating). This phase is necessary for organisms to respond in a timely manner to certain environmental contexts. The intensity and frequency of the CS-US pairings, as well as the salience of the involved stimuli, can all affect the strength of Pavlovian biases (Schultz, 2006). For example, if a loud noise (US) is frequently coupled with a flashing light (CS) eventually the light alone may cause a startle reaction (CR) (Domjan, 2005).

Extinction or inhibition of Pavlovian biases occur when the conditioned stimuli (CS) is no longer paired with the unconditioned stimuli (US) and over time the conditioned response (CR) decreases. Theories have suggested that extinction is not merely a process of forgetting the acquired and learned associations, but the formation of a new memory that is actively competing with the old, conditioned memory (Bouton et al., 2004). The diminished or eliminated response may return in a different context or after a delay due to context-dependent effects caused by the aforementioned competition, a process known as ‘spontaneous recovery’ (Bouton & Peck, 1992). The complete eradication of Pavlovian biases is challenging as these biases have shown perseverance when placed in different contexts than the ones they were acquired or eradicated in. This perseverance is the cornerstone of Pavlovian biases as it demonstrates the capacity to which these biases can impact behavior to cause maladaptive consequences (Myers & Davis, 2007).

2.3 Are Pavlovian Biases Inflexible?

Classical conditioning mechanisms rely on non-declarative and procedural learning (Dere et al., 2014; Pakaprot, 2011; D’Souza et al., 2021). Therefore, Pavlovian biases – learned through classical conditioning mechanisms – are similarly known to exhibit rigid behavior. Attending to ancestral needs, these biases have persevered as one of the decision-making mechanisms that are known to be cognitively low-cost and efficient when necessary. Presently, we encounter arguments that propose the existence of flexibility within these biases, especially in certain contexts - leading to a notable gap in the literature that addresses the question directly.

Recently, Sebold and colleagues found that prolonged stress can increase the rigidity of these learning biases (Sebold et al., 2014). This paper suggests that individuals who suffer from traumatic stress which is usually associated with anxiety and avoidance show strengthened Pavlovian biases. Similarly, there are numerous accounts to support Pavlovian bias inflexibility. When interacting with our goal-directed actions, these biases impede our ability to invigorate action associated with a potential punishment and suppress action when anticipating a reward (Guiatrt-Massip et al., 2012). Much foundational work has been conducted over the years that indirectly addresses Pavlovian bias' core characteristic properties as rigid to environmental needs (Ereira et al., 2021; Mkrtchian, et al., 2017). According to research conducted by Huys and colleagues, Pavlovian biases are automatic and enduring, frequently suppressing and overpowering goal-directed behavior, for example, when an individual is faced with a dangerous situation, such as encountering a wild animal on a trail, it could lead them to jump back with fear or start running regardless of what the rationally best action could have been in the situation (Huys et al., 2017). Therefore, there is a plethora of research supporting the intrinsic rigidity of these biases.

Recently, scholars have started presenting a conflicting argument with some suggesting that perhaps these biases can demonstrate flexibility when necessary. Current work has provided evidence that developmental stage (Raab & Hartley, 2020), motivational value (Dayan et al., 2006), environmental contexts (Cavanaugh et al., 2013) and training (Fleming et al., 2023) can adjust the flexibility of these learning biases. Specifically in 2014, Bouton revealed that these biases can be reduced or even diminished in particular contexts but might show continued perseverance in other contexts, showing contrasts in their behavior (Bouton et al., 2014). In 2020, Raab & Hartley found variability in the interaction between Pavlovian biases and goal-directed learning in populations ranging between children and adults. They found that adolescents exhibit less susceptibility to Pavlovian biases when in conflict with goal-directed actions (Raab & Hartley, 2020). Research has also suggested the Pavlovian biases can capture the motivational value of the outcome presented and employ top-down mechanisms to optimize actions (Dayan et al., 2006). Lastly, training has been proven to decrease or diminish the robustness of these learning biases (Flemming et al., 2023). However, there is a notable gap in research providing evidence to which extent these biases are rigid or flexible.

3.0 Instrumental learning

The second important system which interacts with Pavlovian biases in making decisions is known as instrumental learning. Both Pavlovian biases and instrumental systems work together to determine the final course of action taken by an individual (Dickinson et al., 1995). Similar to Pavlovian biases association with classical conditioning mechanisms, goal-directed instrumental learning also stems from operant conditioning which operates on an action-outcome mechanism (Skinner, 1953). Specifically, within operant conditioning, the response related to the action demonstrated by an individual can either increase or decrease the likelihood of that action occurring again in the future. The famous Skinner operant box experiment demonstrates how pigeons were able to learn how to press a lever to obtain reward (food). In this instance, the pigeon pecking the lever represents the operant response, while the sight of the lever serves as the operant stimulus. When the lever is pecked by the pigeon, food is presented in the form of positive reinforcement therefore over time this learned action-outcome (A-O) association makes it more likely that the pigeon will repeat the behavior. Human learning and decision-making is a complex phenomenon that is largely informed by the two learning systems: Pavlovian bias and instrumental learning. Within instrumental learning there are two additional mechanisms that can be distinguished through their core properties, goal-directed and habitual learning (Shanks, 1993; Balleine et al., 1998; Niv et al., 2007).

3.1 Goal-Directed and Habitual Instrumental Learning

There are two types of instrumental learning: goal-directed and habitual instrumental learning. As the name suggests, the former type of learning occurs when the individual consciously chooses an action, fully aware of the consequences (Balleine & O'Doherty, 2010). This type of learning is highly flexible as it allows individuals to adjust their behaviors to attain desirable outcomes and avoid negative ones (Dickinson et al., 1985). Goal-directed instrumental learning is employed during intellectually analytical tasks such as problem-solving and in situations where individuals have the ability and cognitive freedom to choose the best option to optimize their behavior. For example, when presented with a difficult choice, an individual might consider several factors, weigh the benefits and disadvantages of each, and decide on the course of action that best accomplishes their goals (Eryilmaz et al., 2017). Additionally, action-outcome association awareness is one of the hallmarks of goal-directed instrumental learning system. The basis of instrumental learning originates from an individual's ability to use the feedback from their particular action and adjust as necessary to obtain rewards and avoid penalties. For example, a guitarist might adjust their notes with every sound output to ensure they reach their desired musical composition for a song. Goal-directed instrumental learning has two key features that distinguish it from habitual instrumental learning: flexibility, and higher cognitive processing (Eryilmaz et al., 2017; Chen et al., 2022).

Habitual instrumental learning occurs when responses are made in the presence of a certain stimuli and the agent is then either rewarded or punished. Although habitual learning such as Pavlovian bias is based on the learned association between the stimulus (S) and the response (R), under instrumental learning it sustains its action (A) to outcome (O) association (Lee et al., 2021). Though categorized under habitual mechanisms, this type of instrumental learning offers individuals autonomy to choose a specific response unlike Pavlovian biases where responses are inflexible (Dickinson et al., 2018). Within habitual instrumental learning, participants may be aware of the outcomes following the response, but they continue to perform specific behaviors regardless of the outcome's value (Steding et al., 2019). Rigidity, or the resistance to change that accompanies a habitual behavior can be desirable in surroundings that are stable and need consistent responses, but it can also be maladaptive in situations that are changing and call for flexibility (Wood & Rünger, 2016). Another characteristic that distinguishes habitual learning is the reduced cognitive load. Automatic and low-cost, habitual instrumental learning is known to alleviate cognitive load as it does not involve high analytical decision-making. This occurs because the actions in habitual goal-directed mechanisms occur impulsively, even if the reaction is not optimal to the given situation (Graybiel, 2008). This can be seen in a modern-day example where an individual might grab their phone immediately when it chimes or buzzes in anticipation of a reward in the midst of a study session.

The interaction between the two types of instrumental learning is quite complex. Research suggests that when individuals are put in a new environment with uncertain outcomes associated with the actions employed, goal-directed instrumental learning is utilized, however over time with enough practice individuals start relying on inflexible habitual instrumental learning to reduce their cognitive processing load which can allow space for more novel information (Dolan et al., 2013; Balleine et al., 1998). In the present research, we examined early stages of learning in order to primarily recruit goal-directed instrumental learning.

4.0 Pavlovian Bias and Instrumental Learning Interaction

The interaction between automatic and flexible learning plays a crucial role in decision making. Their interaction can be synergistic when they are in alignment but often can incur complications in behavior when the inflexible Pavlovian bias is not in agreement with the goal-directed action. Pavlovian biases are known to overpower goal-directed actions when they are incongruent, due to their rigid nature. Huys and colleagues found behaviors such as approach and withdrawal were employed differently when combined with a Pavlovian stimulus that corresponds to reward (Huys et al., 2007). For example, they reported that when the outcome was positive (reward), it led to action, however the same outcome presented challenges when associated with withdrawal or suppression of action. Hershberger’s chicks are a prime example of an animal’s inability to realize that walking away from food would lead them to acquire it

(Hershberger, 1986). When goal-directed learning and Pavlovian biases are present in the same task, as in the one used in the current study, we hypothesized that we would observe flexibility differences between the goal-directed learning and Pavlovian biases. Pavlovian biases are typically less flexible and do not adapt to shifting circumstances. Goal-directed learning, on the other hand, is distinguished by its flexibility, which enables people to alter their behavior in response to modifications in action-outcome contingencies (Dolan & Dayan, 2013). Studying this relationship of how these two systems interact can elucidate new avenues for research and clinical interventions that can promote optimized learning and habits.

4.1 Methods for Measuring PIT

Over the years, researchers have developed several paradigms to explore the complex relationship between goal-directed instrumental learning and Pavlovian biases (Talmi et al., 2008; Raab & Hartley, 2020; Nadler et al., 2011). One of the most prevalent tasks utilized within non-human animal research is the Pavlovian to Instrumental Transfer Task (PIT). In this task animals are taught to learn the association between a presented stimuli and the reward that it produces over time (Pavlovian bias) and examine how these associations can either impede or facilitate their behavior in the task that presents them with rewards or penalties (instrumental learning). Specifically, there are three phases in a PIT task, namely: instrumental conditioning phase, Pavlovian conditioning phase, and lastly, the transfer phase (Costa et al., 2020; Mahlberg et al., 2021; Degni et al., 2022).

There are a variety of tasks that can be utilized with human participants to assess the interactions between Pavlovian bias and goal-directed instrumental learning. Some of these tasks include: the Saccadic Response Task (SRT) with reward and punishment cues (Rothkirch et al., 2012); the Implicit Association Task (Boschen et al., 2006); the Affective Priming Task

(Hermans, 2002); and the Automatic Approach Avoidance Task (Zech et al., 2020). However, the most commonly used tasks are a human version of the aforementioned PIT and the Pavlovian go/no-go task, used in the present study. PIT consists of two training stages followed by a transfer stage. The first stage in PIT is instrumental training which involves participants ability to learn actions (pressing button or moving joystick) associated with specific outcomes such as reward or avoiding penalties. This is followed by Pavlovian training where participants are taught which stimuli corresponds to which outcome (reward/avoid punishment). The last stage consists of Pavlovian and instrumental interaction where the Pavlovian cues taught earlier are used to facilitate or impede the instrumental actions participants had learned, this is called the Pavlovian-instrumental transfer stage. When positive Pavlovian cues are present, PIT is characterized as an increase in responding or response vigor, and when negative Pavlovian cues are present, a reduction in responding or response vigor (Peng et al., 2022).

There are additional tasks that can be combined with PIT to assess inhibition or facilitation resulting from Pavlovian biases. The outcome devaluation task modifies the reward's value and can be used to determine whether behavior is under goal-directed or habitual control. After the desirability of the reward is changed (devalued) the organism will stop responding if the behavior is under goal-directed control but will continue to respond if it is under habitual control (Balleine & Dickinson, 1998).

An alternative way to measure PIT, and the one used in the current study, is the Pavlovian go/no-go task. The go/no-go task was explicitly designed to evaluate how Pavlovian valence (i.e. reward or punishment) interacts with instrumental action (i.e. going or not going) (Guitart-Masip et al., 2014). The task incentivizes participants to engage in instrumental learning to acquire as many points as possible; however, their choices can be facilitated or impaired based on the Pavlovian bias (Ereira et al., 2021). Therefore, this task is ideal to examine the interactions between the two learning systems as well as the extent of the biases present. The two actions (go vs no-go) within the task are orthogonalized and independent of the outcome that they incur. The goal outcomes within this paradigm are winning and avoiding loss. Therefore, the cross between the actions and the outcomes results in four different conditions namely: Go to Win (GW), Go to Avoid Losing (GAL), No Go to Win (NGW), No Go to Avoid Losing (NGAL). Broadly, within this task participants exhibit two major biases: motoric action of 'go' when presented with a reward due to automatic invigoration and secondly, suppressing action of 'no-go' when a penalty is associated with a particular action. The general trend within this go/no-go task is participants over reliance on action or 'going' than inaction (Ereira et al., 2021). This is suspected to align with an individual's idea to have a choice and having a sense of control when choosing action over inaction. This finding is also supported by the trend where participants learning the action-outcome association overtime perform significantly better on the 'go' trials than the 'no-go' trials – specifically when their actions are in contingency with rewards and inaction with avoiding penalties (Guitart-Masip et al., 2012).

5.0 Neural Underpinnings of Pavlovian Biases and Goal-Directed Instrumental Learning

The emotional processing region, amygdala and the reward processing neural system ventral striatum are both involved in the process of carrying out Pavlovian biases (Bray et al., 2008). Conversely, the prefrontal cortex and dorsomedial striatum (DMS) are involved in regulating and employing goal-directed decision making (Balleine & O'Doherty, 2010). It becomes imperative that we study the neural distinction between instrumental learning and Pavlovian biases to understand them extensively and account for differences within their biological makeup, especially in how they support human learning.

5.1 Neural Mechanisms for Pavlovian Biases

Aversive Pavlovian conditioning relies on the amygdala as shown by studies examining fear conditioning in healthy agents in which the amygdala has been temporarily inactivated and in patients with amygdala lesions (Muller et al., 1997; Helmstetter & Bellgowan 1994). Within the amygdala, the Basolateral (BLA) and central nuclei (CeA) are essential components activated during Pavlovian conditioning. The establishment of connections between emotionally meaningful outcomes or valence (such rewards or punishments) and neutral stimuli are facilitated by the BLA. Conversely, these learned emotional reactions cannot be expressed without the CeA (LeDoux, 2000). Additionally, to process motivational aspects of reward related Pavlovian biases, the ventral striatum, including the nucleus accumbens (NAc) is necessary. Prefrontal cortex and amygdala inputs are combined via the NAc to appropriately respond to the positive motivational aspects of the conditioned stimuli (Ishikawa et al., 2008). Another key role is played by the VTA or ventral tegmental area that supplements the NAc with dopaminergic input, helping it sustain the Pavlovian association by increasing motivation to aquire the reward or avoid punishment (Nestler & Carlezon, 2006; Berridge & Robinson, 1998). The amygdala, NAc and prefrontal cortex are the main brain regions that receive the dopaminergic signal from the VTA. The general anticipation of reward releases dopamine from the VTA which reinforces and sustains Pavlovian learning, similarly with unexpected rewards, the link between the stimulus and outcome becomes more strongly associated when the dopamine input is received from the VTA to these major brain regions (Schultz, 1998).

5.2 Striatal Mechanisms for Goal-Directed and Habitual Instrumental Learning

The two systems within instrumental learning recruit separate striatal mechanisms. Specifically, the dorsal medial striatum (DMS) plays a crucial role in goal-directed actions while the dorsal lateral striatum (DLS) is involved in the processing of habitual behaviors (Gremel & Costa, 2013). Research shows that damage to the DMS can cause habitual behaviors to strengthen and a lack of goal-oriented actions. Conversely, damage to DLS can disrupt habitual behaviors, allowing goal-directed actions to take precedence. This supports the evidence of dynamic interplay between automatic and flexible learning mechanisms in the brain. In case of the shift that could be caused due to lesions or higher cognitive control, the lateral orbitofrontal cortex (OFC) plays a significant role in this shift. Additionally, the basal ganglia support the neural shift from goal-directed to more habitual processes within the brain (Gremel & Costa, 2013).

5.3 Cortical Mechanisms for Goal-Directed Instrumental Learning

The prefrontal cortex (PFC) is responsible for a large majority of planning and decision-making within our brain, including goal-directed learning. This brain region assists in processing the outcome associated with the action being performed and has the ability to modify action when a better outcome is available. The PFC integrates information and directs goal-directed activities by receiving input from many sensory and limbic regions as well (Miller & Cohen, 2001). The basal ganglia, specifically the dorsomedial striatum (DMS) is responsible to integrate information from the PFC about action-outcome associations. An additional system that assists our goal-directed actions is the hippocampus which facilitates the creation of action-outcome linkages by offering insights from previous learned events and contextual information. It aids in encoding the context of the environment in which behaviors are carried out, allowing the organism to remember previous encounters and modify behavior accordingly (Eichenbaum, 2000; Wimmer et al., 2012). Previous research suggests that the striatum and prefrontal cortex interact with the hippocampus to assist in flexible, goal-directed behavior (Graybiel, 2008).

Finally, there are neural systems that help to arbitrate between goal-directed and habitual learning. The two systems use independent neural systems with DMS involved in goal-directed activities and DLS for habitual (Gremel & Costa, 2013). When our behavior changes from habitual to goal-directed, the OFC or orbitofrontal cortex is the main brain region necessary for this shift. Conversely, the basal ganglia is the key brain region that allows the shift to be made from more goal-directed to habitual processes (Gremel et al., 2013; Yin et al., 2006). However, the inferior lateral PFC arbitrates which mechanism to employ and when to do so.

5.4 Neural Mechanisms for Interactions Between Goal-Directed Instrumental Learning and Pavlovian Biases

Switching between our flexible actions to automatic ones and vice versa is critical to sustain human learning and decision-making processes. Research has suggested the crucial role of inferior lateral prefrontal cortex (ilPFC) and frontopolar cortex as the modulator between our goal-directed actions and Pavlovian behaviors (Kim et al., 2019; Gruner et al., 2015; Lee et al., 2014; Bogdanov et al., 2018; Kim et al., 2018). In 2015, Lee and colleagues found that ilPFC and right medial frontopolar cortex consider predictability of outcomes and degree of reliability between the two learning systems and choose the one that optimizes the decision (Lee et al., 2014). In ilPFC, the individual reliability signal (signaling which mechanism has no conflict of interest in the given situation) and the strongest of the two systems is accounted for, however in inferior lateral frontopolar cortex, only the reliability signal is considered. Lee and colleagues also stated that the connection strength between the ventromedial prefrontal cortex (vmPFC) – which encodes the integration of value – and putamen changes depending on the decision-making mechanism being employed. Specifically, in a situation favoring goal-directed decision-making actions, the strength of these aforementioned valuation systems is weakened through the arbitrator or the ilPFC. Therefore, the inferior lateral prefrontal cortex has been suggested to evaluate the reliability of the two learning systems in a given situation and modulate the strength of habitual behavior over our flexible goal-directed actions (Bogdanov et al., 2018; Kim et al., 2018).

6.0 Maladaptive Effects of Pavlovian Biases in Disorders

Pavlovian biases can facilitate complex cognitive processes by allowing us to rely on those conditioned responses in situations of low predictability and control and in unfamiliar situations, however they can also create impediments to our goal-directed learning and solidify our reliance on maladaptive behaviors (Ereira et al., 2021). These undesirable behaviors range from binge eating to substance use disorders.

6.1 Substance Use Disorders (SUDs)

Substance use disorders are defined as maladaptive use of drugs despite severe adverse consequences (Bush & Lipari, 2016; Brady et al., 1999; Han et al., 2009, Rosenthal et al., 2022). Over the years, Pavlovian biases have been strongly associated with eliciting and sustaining Substance Use Disorders (SUDs). Research has found that one of the primary mechanisms through which Pavlovian biases affect SUDs is Pavlovian-to-Instrumental Transfer (PIT) (Cartoni et al., 2016; Everitt et al., 2001; Glasner et al., 2005). PIT is a strong candidate to explain how habitual behaviors are formed over time – especially through heightened cue-reactivity; a salient mechanism in the development of addictions and SUDs (Gusburrow et al., 2014; Hansson et al., 2017; Sebold et al., 2021). Cue-reactivity is defined as cognitive or physical responses (conditioned response) to stimuli associated with drugs – conditioned stimuli (Garland et al., 2012; Childress, 2016). These conditioned stimuli can influence our goal-directed behaviors, leading to an overreliance on Pavlovian biases when interacting with conditioned cues. This process is particularly relevant within individuals with SUDs as environmental cues can prompt intense drug cravings and drug seeking behaviors even after long periods of sobriety or temperance, leading to relapse (Vafaie & Kober, 2022; Ebrahimi et al., 2023).

6.2 Obsessive Compulsive Disorder (OCD)

OCD or obsessive-compulsive disorder can be characterized as one having obsessive thoughts and compulsive behaviors (Stein, 2002). These thoughts are conditioned responses to alleviate stress or anxiety (Raposo-Lima & Morgado, 2020). The compulsive behaviors observed in OCD may be associated with enhanced Pavlovian responses to distressing or frightening stimuli, as people come to link particular acts with momentary relief from anxiety (Gillan et al., 2011). For example, washing hands excessively after touching a doorknob (CR) to reduce the fear or anxiety of spreading contamination (CS). Studies reveal that the cortico-striato-thalamo-cortical (CSTC) loop, an important circuit for habit formation and Pavlovian training, is abnormally active in OCD patients (Menzies et al., 2008). According to Menzies and colleagues, dysregulation in these circuits results in increased Pavlovian biases and maladaptive learning, which feed compulsive behaviors. Research suggests that in OCD, instrumental learning can dominate which is able to potentially override Pavlovian biases (Peng et al., 2022). This suggests that while Pavlovian associations—such as responding to an aversive cue with anxiety or fear—are a major contributing element to OCD, they are not entirely inflexible. Rather, goal-directed, habit-forming processes have the ability to moderate or even suppress them, especially when these processes become dominant.

6.3 Anxiety & Depression

Heightened sensitivities to Pavlovian cues within the environment can lead to excessive fear and avoidance and ultimately to development of anxiety disorders (Hofmann & Hay, 2018). Specifically, fear conditioning towards day-to-day environmental stimuli is putatively known as the root cause of anxiety (Modecki et al., 2022) meanwhile, individuals with depression have been known to demonstrate sustained conditioned responses to aversive cues (Wiggert et al., 2016). Past research states that anxiety and depression are informed through multiple brain regions such as: amygdala responsible for fear conditioning (Martin et al., 2009), insula and anterior cingulate cortex regulating awareness and emotional processing (Martin et al., 2009), prefrontal cortex responsible for executive functioning and optimal decision-making (Bouras et al., 2023), hypothalamus modulating fear conditioning (Fischer, 2021), and hippocampus responsible to hold aversive memories (Chen & Etkin, 2013), to name a few. If dysregulated, these brain regions can support maladaptive behaviors such as anxiety and depression. Depression can be associated with higher sensitivity towards negative events while diminished sensitivity towards rewards, usually caused by negative reinforcement through Pavlovian biases. This can result in maladaptive behaviors like rumination and anhedonia. According to Whitton et al. (2015), these biases have the power to reinforce unfavorable mental patterns and obsessive activities meant to lessen discomfort (Whitton et al., 2015).

6.4 Eating Disorders

Bulimia and anorexia nervosa are two of the major eating disorders that are characterized by harmful eating patterns as well as distorted body image issues (Kaye et al., 2009). Research suggests that an imbalance within the reward circuitry can play a significant role in evoking and maintaining eating disorders – especially if our goal-directed actions become overpowered by

Pavlovian biases (Avena & Bocarsly, 2011). These biases attach valence to particular food or body related cues eliciting feelings of comfort or relief when in a stressful situation (Park et al., 2014). In 2015, Voon and colleagues addressed the role of Pavlovian biases in people who had eating disorders. The authors reported that enhanced Pavlovian biases are strongly connected with cues related to food which can lead to destructive eating habits. The obsessive component found in the behavior of these individuals can be attributed to added rigidity to these conditioned responses (Voon et al., 2015).

6.5 Chronic Pain

One of the recent advancements in research is exploring Pavlovian bias's role within chronic pain patients. Research defines chronic pain as any pain that persists or reoccurs for more than three months (Rolf-Detlef, 2018). Chronic pain can include lasting migraines, lower back pain, fibromyalgia (widespread muscle pain) and even arthritis. Similar to acquisition and maintenance of aforementioned disorders, chronic pain can be exaggerated by developing heightened sensitivities to conditioned stimuli especially those associated with avoidance of painful stimuli (Harvie et al., 2017). Harvie and colleagues suggest that overgeneralization of pain can lead to robustness of chronic pain, especially within older populations. Additional research supports this idea by suggesting that behaviors that promote induction of pain might occur from reinforcement, specifically from avoidance of negative events (Fordyce et al., 1976). With a majority of older population affected by chronic pain, it is imperative we study the role Pavlovian biases when they go awry and their role in acquisition and maintenance of cognitive disorders.

6.6 Therapies addressing Pavlovian biases

Therapies specifically effective for maladaptive behaviors require an understanding of the interaction between goal-directed learning and Pavlovian biases. For instance, addiction therapies must address both the cognitive, goal-oriented components of instrumental learning as well as the rigid, stimulus-driven reactions of Pavlovian biases, since both goal-directed and Pavlovian processes may be at work. Interventions can be more thorough and successful in encouraging flexible behavior change by focusing on both processes (Everitt & Robbins, 2005). Natural interventions such as Mindfulness Based Intervention (MBI) combined with cognitive behavioral therapies (CBT) are capable of addressing both the goal-directed and cognitive as well as inflexible aspects of maladaptive behaviors, making interventions more comprehensive and effective (Zhang et al., 2021). The aim of the aforementioned interventions is to strengthen cognitive flexibility to allow goal-directed actions to take control over rigid biases (Lange et al., 2017).

7.0 Using Reversal Paradigms to Assess Declarative Mechanisms and Flexibility

Cognitive flexibility is essential for learning and development of healthy cognitive functions (Roy & Dugal, 1998; Rudnik et al., 2019; Dajani & Uddin, 2015). Reversal paradigms have successfully been used to assess a wide range of types of behavioral flexibility (Harris et al., 2020). These paradigms often have two phases: a training phase where participants are taught the associations that a stimulus has with action and outcome; and a reversal phase where the previously learned contingencies are switched by either changing the outcome associated with the stimuli or the action (Weiss et al., 2021). In reversal paradigms, once a participant has hit threshold for the learning criteria by correctly answering certain number of correct answers (Dalton et al., 2014) or completing a fixed block of trials (Farashahi et al., 2017), the associations between stimuli and reward are switched. After reversal, previous responses are now unable to provide individual with the same rewards, however, they remain frequently chosen due to cognitive rigidity. Therefore, these paradigms require cognitive flexibility as one has to suppress the previously learned responses and start acquiring new associations. In 2021, Weiss and colleagues examined cognitive flexibility within children and adolescents via reversal learning (Weiss et al., 2021). In this study, they found that children performed worse than adolescents when subjected to reversal of task contingencies. Specifically, children performed worse overall and made more regressive errors, meaning that they kept using the same rules even after reversal occurred, regardless of if those rules were not right anymore. Additionally, in 2006, Remijnse and reversal learning has been shown to be impaired in participants diagnosed with Obsessive-Compulsive Disorder (OCD), highlighting the difficulty an individual could face, especially in anticipation of reward, when the interaction between Pavlovian biases and/or instrumental learning becomes complicated (Remijnse et al., 2006; Gold & Shadlen, 2007). This can result in more inflexible and inappropriate reactions when circumstances change.

Reversal paradigms are appropriate for assessing the flexibility of both Pavlovian biases on decision making and instrumental learning. This paradigm successfully captures whether the knowledge participants acquire during the task is goal-directed or skews towards habitual learning: if the participants can flexibly shift to the new associations, that is evidence that performance is mediated by goal-directed mechanisms, whereas if performance declines after reversal then it is mediated by rigid learning mechanisms. In our study, we aim to compare reversal of Pavlovian learning biases – which is expected to decrease accuracy in participants performance – with goal-directed instrumental learning reversal – which will not be affected by the reversal significantly due to its inherent declarative nature.

8.0 Overview of Experiment 1 and Experiment 2:

There are two main goals of this study. One is to gain a deeper understanding of how Pavlovian biases interact with the goal-directed instrumental learning system. Specifically, we want to examine the impairment or facilitation that occurs due to the coupling of valence and action within the goal-directed instrumental learning system. Secondly, this study explores Pavlovian biases rigidity in context of reversal learning. We hypothesize that since Pavlovian biases rely on inflexible or procedural learning, participants will encounter challenges post reversal due to resistance exhibited by these biases. This study consists of two experiments, each of which consists of three subparts (a, b, c) namely, baseline performance for the Pavlovian go/no-go task, reversal of goal-directed instrumental learning and lastly, reversal of Pavlovian bias. Experiment 1 preliminary results were affected by noise and participants' inability to learn the task in the pre-reversal phase. Therefore, in Experiment 2 we introduced four main changes to Experiment 1 paradigm to ensure the task was easier to learn in order for participants to gain adequate proficiency within the pre reversal phase to establish robust findings post reversal. First, we incorporated an extra block to extend the pre-reversal learning phase for participants. Second, we shortened the duration of each trial to ensure that the task is engaging for participants. Third, we changed the probabilistic assignments of ‘win’ or ‘avoid loss’ conditions from 80-20% to 100%, making the task fully deterministic. Lastly, we included a post-task survey to examine participants ability to acquire contingencies of action (go/no-go) and valence (win/avoid loss) explicitly or implicitly.

EXPERIMENT 1

For our first experiment we had two objectives: examine how Pavlovian biases and goal-directed instrumental learning interact and assess the rigidity of Pavlovian biases through a reversal paradigm. Experiment 1 had three subparts that were collected sequentially: baseline task without reversal (1a), reversal of goal-directed instrumental learning (1b) and reversal of Pavlovian biases (1c). Firstly, in line with previous research, we anticipated the presence of Pavlovian biases within participants task performance often overpowering goal-directed instrumental learning during incongruent trial types (NGW and GAL). Secondly, we hypothesized that after reversal, Pavlovian biases will show rigidity and result in compromised performance by participants, but when goal-directed instrumental learning is reversed, we will see a more flexible performance due to their reliance on declarative and explicit learning.

Methods

Participants

Participants were recruited from PSY100 and PSY250 research pool. Recruitment was conducted after conducting power analysis that reported that for a target significance level of 0.05, and a desired power of 0.8, the study would be sufficiently powered for a target effect size of .49 if we recruited 30 participants per condition. Therefore, we recruited approximately 30 participants for each of the three conditions: baseline (no reversal), reversal of goal-directed instrumental learning, and reversal of Pavlovian Biases. In total we recruited approximately 90 participants for Experiments 1. The participants ranged in age between 18 and 24 years (*M* = 19.50; *SD* = 1.50) with 31 males, 55 females in the dataset with 2 unreported gender identities.

Behavioral Task Design

We used the version of the Pavlovian go/no-go task developed by Raab and Hartley (2020) illustrated in Figure 1. In the task participants were shown four different stimuli (gamified robots) which were associated with one of the four different action outcome contingencies. The contingencies involved two instrumental responses (go/no-go) paired with one of the two Pavlovian valence (win/loss), resulting in four conditions as follows: Go to Win (GW), Go to Avoid Losing (GAL), No Go to Win (NGW) and No Go to Avoid Loss (NGAL). Within the task participants were told to maximize the rewards by avoiding loss and obtaining rewards. The task consisted of 3 blocks with 60 trials each (180 trials total) where the four contingencies were distributed pseudo-randomly within each block. Each trial type was presented 45 times across the 3 blocks. Before beginning the task, participants practiced pressing the button to indicate "Go" and not pressing the button to indicate "No-Go”. In order to evaluate which robots would provide (or take) a ticket for one action but not for the other, participants tried pressing and not pressing the spacebar for each type of robot, which was a part of their pretraining. After each block, participants were given a break that included their score for the particular block. Each trial began with a display of one of the four robots (cue) for 1000ms which was followed by a fixation point for 1000ms and then the target was shown where participants could either press or withhold response. The response window was about 1500ms, therefore if participants chose to press the button of the robot (target) they would be presented with another fixation point for 1500ms.

Finally, participants were shown the outcome for 3000ms where they either received a golden ticket (win), horizontal bar (avoid loss) or broken ticket (loss).

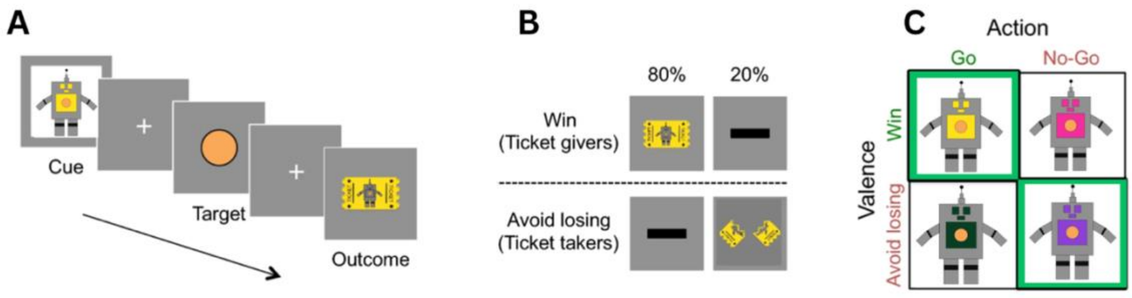


Figure 1:Task adapted from modified version of go/no-go (Raab & Hartley, 2020). A. In one trial, participants first saw the cue, then target and finally, the outcome. The cue could be one of the four differently colored robots, however the target (orange button) remained consistent with each trial. Outcomes were a golden ticket (reward), broken ticket (loss) or horizontal bar (avoid loss). B. Represents the probability of the contingencies when participants were correct: 80% of the time, when winning they would see the golden ticket however 20% of the time they would be presented with the horizontal bar (meaning they had avoided loss), similarly, when correctly responding to an avoid loss contingency, 80% of the time participants would see avoid loss (horizontal bar) while 20% of the time they would be shown a broken ticket (loss). C. The four differently colored robots representing four different contingencies.

Baseline Task & Reversal Conditions

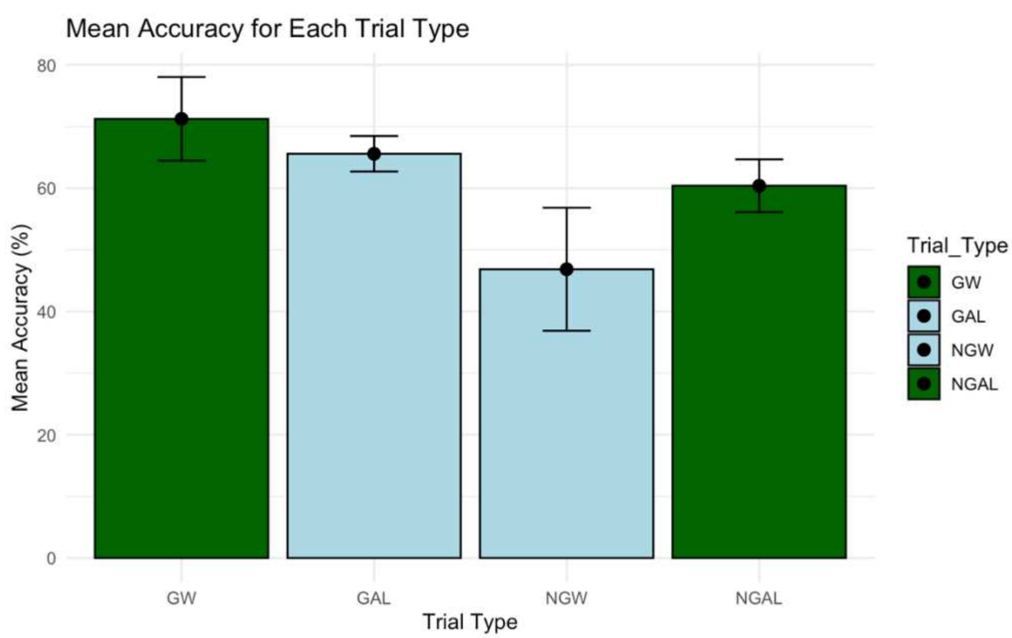
Experiment 1 included three different experiments, two of which used a task reversal paradigm to examine the effects of reversal of either goal-directed instrumental responses (1b), or Pavlovian biases (1c) on learning. We compared the results with a first experiment that included no reversal (1a). There was a total of 3 blocks in each of the Experiment 1 versions. Participants initially performed the standard task described above, for the initial block out of the three blocks, and then performed blocks 2 and 3 under the reversal condition for Experiments 1b and 1c. All three blocks, regardless of reversed conditions or not, had the same layout and structure within the paradigm to ensure consistency. In the beginning of the experiment participants were instructed to remain attentive to the rewards associated with each robot, but no specific instructions on reversal were given to the participants. This ensured that they would learn about the reversal through trial and error.

Reversal was implemented in two ways, reversal of goal-directed instrumental learning and reversal of Pavlovian biases. To test the effect of reversal on goal-directed instrumental learning we switched the action associated with each robot from ‘go’ to ‘no-go’ and vice versa. Therefore, a robot that was previously programmed to present a reward when an action was completed by the participant, now gave a reward when the participant withheld their response. Conversely, to test the effect of reversal on Pavlovian biases, we altered the outcomes associated with the four robots, switching from ‘win’ to ‘avoid loss’ and vice versa. This alteration meant the response previously performed as ‘Go to Win’ now became ‘Go to Avoid Losing’ and so on. We provided no instructions on the specificities of reversal albeit instructed participants to carefully observe and learn the contingencies as they change.

Results

Experiment 1a: Baseline Task (Raab & Hartley Replication with College Students)

The initial experiment we conducted was the baseline task with no reversal. The purpose of the baseline was two folds. First, it served to replicate Raab & Hartley (2020) with adult participants in order to ensure that the basic task findings were replicable and appropriate for use in the reversal studies. Second, it functioned as the baseline against which we could compare the two reversal conditions. We first examined accuracy in the four conditions across all blocks, as shown in Figure 2. Superior performance of trial types where Pavlovian biases facilitated instrumental learning (GW & NGAL) can be seen, accompanied by lower performance in conditions in which the responses supported by two learning systems are incongruent (NGW & GAL). This overall pattern is consistent with that found by Raab & Hartley for adult participants, although we found lower mean accuracy across conditions than they did. For statistical comparison of these conditions, see the post hoc tests presented below.



Experiment 1a/Baseline condition: Mean Accuracy of the four trial types across three blocks.

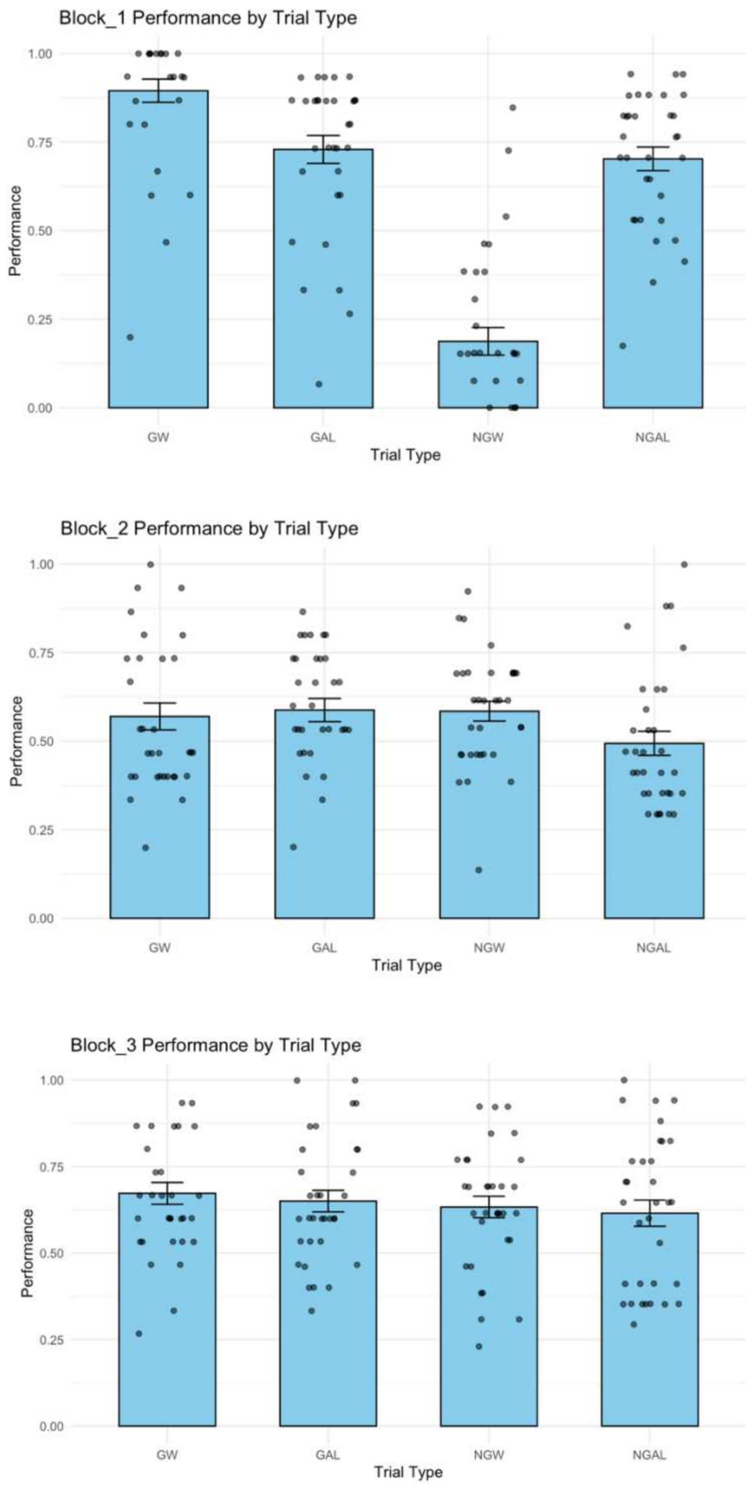


*Figure 3: Experiment 1a/Baseline Experiment: Mean average accuracy in each block. **GW (Go to Win), GAL (Go to Avoid Losing), NGAL (No Go to Avoid Losing), NGW (No Go to Win); p = 0.05**

To visually depict the effect of both block and trial type, and their interaction we plotted a mean performance separately over the three blocks. As shown in Figure 3, this revealed that strong differences between the trial types were limited to the first block. Mean accuracy in blocks 2 and 3 were very similar across conditions. This result was surprising as to our knowledge no previous study has found effects only during early learning, and we found clear trial type differences in blocks 2 and 3 in our reversal conditions. Despite various efforts, we were unable to determine a reason for these observed results or potentially attribute them to a problem with the program during data acquisition. In Experiment 2 we made several adjustments to the procedure and examined whether there is a similar effect of block or not.

The effects of the block and trial type (GW, GAL, NGW, NGAL) on accuracy, as well as the effect of their interaction, were investigated using a 4 (Trial types) x 3 (blocks) two-way ANOVA. Both block ($F(2,5568) = 11.19$, $p < .001$) and trial type ($F(3,5568) = 167.30$, $p < .001$) showed substantial main effects in the analysis, suggesting that both factors independently affect accuracy. Furthermore, the interaction effect between trial type and block was significant ($F(6,5568) = 260.25$, $p < .001$), indicating that the influence of block on accuracy differed based on the particular trial type and vice versa.

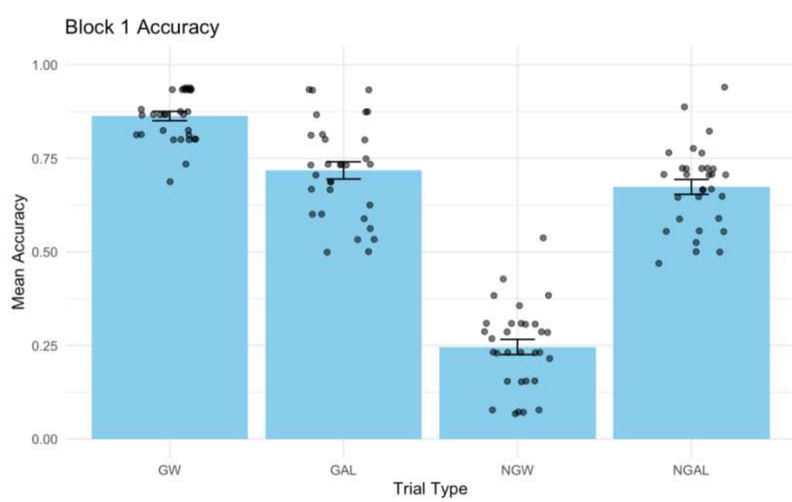
In addition to the two-way ANOVA, we also conducted the Scheffé post-hoc tests to evaluate pairwise comparisons on the effect of trail types on performance. Throughout the task, NGW had significantly lower accuracy than GAL ($p < 0.001$). Similarly, NGW and GW ($p < 0.001$) were significantly different as well, with GW with higher accuracy rates. Lastly, NGAL and GAL ($p < 0.001$) were also significantly different with NGAL having lower accuracy. No significant differences were found between the GW and GAL conditions ($p = .4258$) or between the NGAL and GAL conditions ($p = .5052$). This finding suggests that the least accurate trial throughout the three blocks was NGW potentially due to the incongruency effect between biases and goal-directed actions.

In line with the research of Raab & Hartley, our findings demonstrate the that we can elicit Pavlovian effects in this paradigm. However, overall accuracy was much lower than in Raab & Hartley and other previous studies using this task, and in some conditions was at or below chance. This suggests that many individuals were not successful at learning all of the task associations. By examining individual participant data points (see Figure 3), it is clear that individuals differed greatly in what they learned; for example, some successfully learned the No-Go to Win (NGW) condition, while others appear to have responded solely based on Pavlovian valence and their associated approach-avoidance biases, consistently selecting ‘go’ for any win stimulus and therefore performing significantly below chance in No-Go to Win (NGW).

As mentioned before, we modified the task conditions for Experiment 2 in order to increase the average accuracy and aid participants to learn the associations better so as to obtain robust post reversal results.

Experiment 1b: Results of Reversal of Instrumental Learning:

For the instrumental learning condition, where participants' goal-directed actions were tested through the reversal of actions (e.g., ‘Go to Avoid Losing’ becoming ‘No Go to Avoid Losing’), we hypothesized that the reliance of this learning on flexible goal-directed mechanisms would allow participants to effectively reverse their behavior. Figure 4 shows overall accuracy for each condition within each block. The two-way ANOVA revealed significant effects of block ($F(2, 5748) = 10.40$, $p < .001$) and trial type ($F(3,5748) = 21.26$, $p < 0.001$). Additionally, we found the interaction between block and trial type to be highly significant as well ($F(6, 5748) = 138.33$, $p < .001$). These findings suggest that participants' performance differed substantially throughout the blocks indicative of their ability to reverse contingencies as the changes were introduced. This flexibility is consistent with our theory that flexible goal-directed instrumental learning can adapt to shifting action-outcome linkages. To further asses the effect of different blocks and trial types on accuracy, Scheffe's post hoc test was conducted. The accuracy between blocks 1 and 2 ($p < 0.001$) and blocks 2 and 3 ($p < 0.001$) differed significantly but there was no discernible change between blocks 1 and 3 ($p = 0.985$). The mean accuracy of block 1 was higher than block 3 (0.639 > 0.579), indicating compromised performance post-reversal. For trial types, we found GW to be significantly different from NGW ($p < 0.001$) with GW consistently reporting higher accuracy throughout the task while trial type GAL and GW ($p = 0.07$) did not differ significantly alongside NGAL and GW ($p = 0.372$). We saw an overarching trend of incongruent trial types (GAL and NGW) significantly reporting lower average accuracy than congruent trial types (NGAL and GW) pre and post reversal.



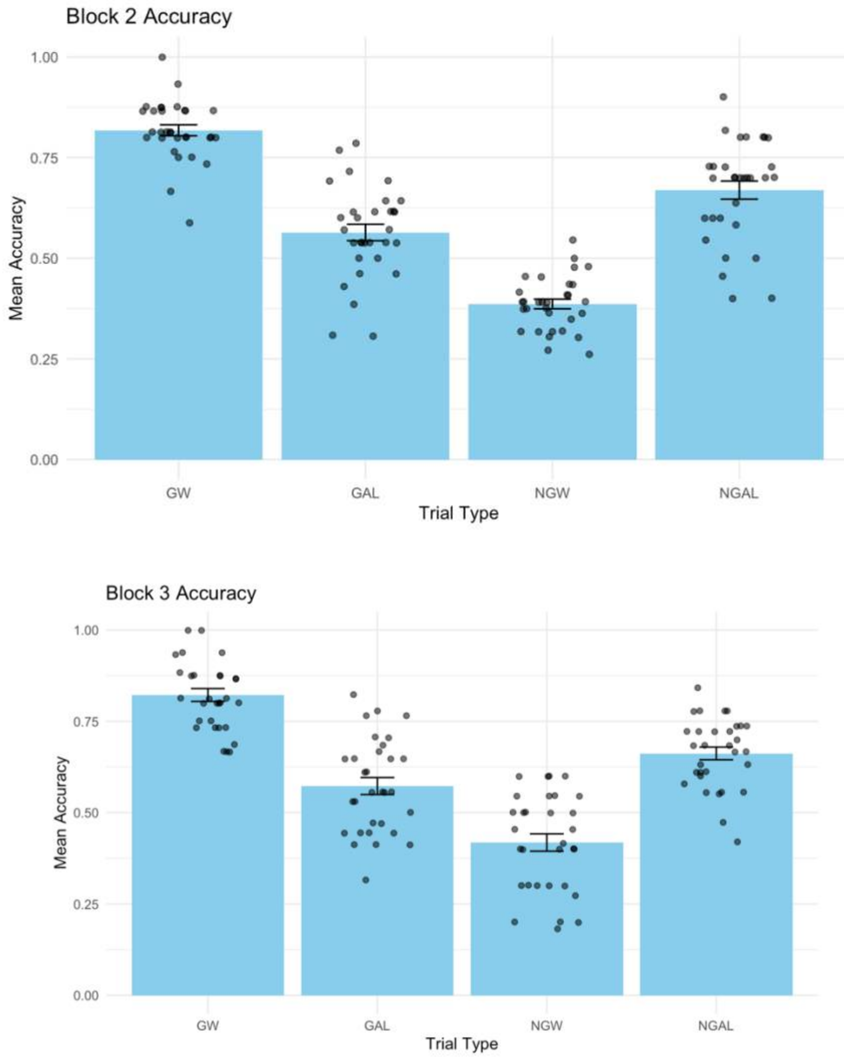


Figure 4: Experiment 1b/Reversal of goal-directed instrumental learning: Average accuracy reported of participants in each block with individual data points. Action was switched in block 3 from 'go to 'no-go' and vice versa.

Focused Examination of Each Reversal Type

To acquire a more detailed visual examination, we plotted trial-by-trial graphs over two blocks (pre and post reversal) with learning curves for each trial type performance. This allowed us to compare the responses to each stimulus (i.e., robot) before and after reversal. We could then assess whether participants learned the initial behavior before reversal and whether they could adjust to the reversal by flexibly adapting their instrumental behavior.

Go to Win changed to No Go to Win

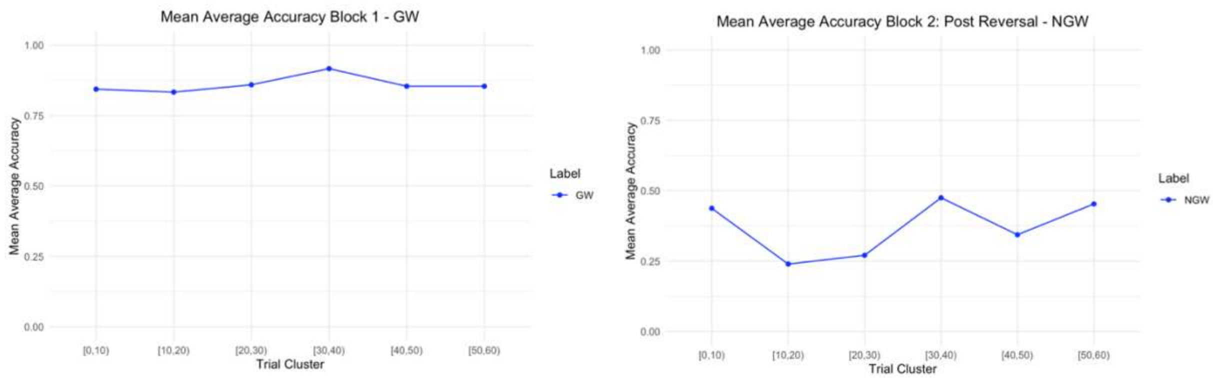


Figure 5

<u>Pre-reversal</u>: ‘Go to Win’ had a high average accuracy throughout the block as participants were able to sustain their performance most likely due to the alignment of both the instrumental learning response of ‘go’ and the Pavlovian bias of ‘win’. The anticipation of a reward became a motivator for the participants to continue to choose action over inaction, relying on their inherent biases of associating ‘go to win’.

<u>Post-reversal</u>: Post reversal the stimulus became ‘No Go to Win’ where the instrumental learning aspect of the stimuli was now switched from ‘go’ to ‘no-go’ while the bias of ‘win’ remained the same. The average accuracy of participants fell below chance suggesting that their ability to associate the no-go response with a reward was compromised. This could have a couple of explanations. First, participants may have been relying on Pavlovian learning across both blocks. After the reversal of instrumental action, participants should have started relying on withholding their response however, since the Pavlovian biases were the same ‘win’ their preferred reaction to the same robot continued to remain as ‘go’ which aligns more with their internal biases of action and reward associations. Secondly, participants may have learned the instrumental response but not have been able to switch their instrumental reversal from ‘go’ to ‘no go’ as it was now associated with a positive outcome, which ultimately resulted in a low average

accuracy. The failure of participants to learn the reversal of the instrumental condition from 'go' to ‘no go’ played a significant role in the decline of their performances.

No Go to Win became Go to Win

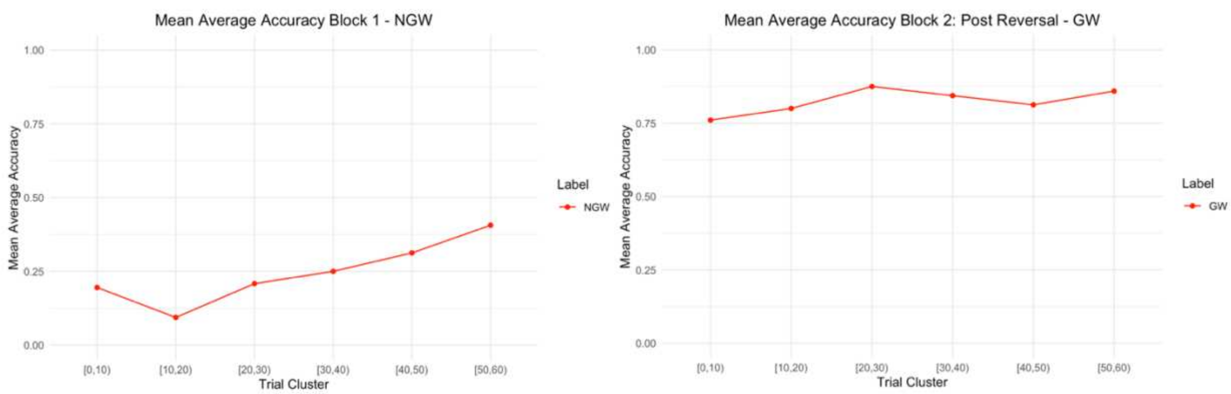


Figure 6

<u>Pre-reversal</u>: No Go to Win (NGW) had a steady increase from the low initial accuracy performance from the participants but never reached chance levels of performance. This suggests that participants were unable combine the incongruent instrumental learning ‘no go’ response and Pavlovian bias ‘win’ together (Figure 6). Because participants had to inhibit a natural inclination to act in the context of a potential win, the instrumental learning process was challenging for them, as seen by the comparatively slow learning rate. Even at the end of the block average performance was still below chance which indicates many participants were unable to learn the association well.

<u>Post-reversal</u>: Post reversal NGW transformed into ‘Go to Win’. Figure 6 shows sustained high accuracy after the instrumental response switched from an incongruent action to a congruent one. Participants' ability to reverse the instrumental condition rapidly suggests that they had already learned the Pavlovian value of the robot (win) in pre-reversal condition (which would explain their low below-chance accuracy pre-reversal). We can attribute the changes we saw in the

participants performance to flexible use of instrumental learning as well. Pre-reversal participants started to learn the associations of inaction with reward, as seen by their steady increase in accuracy. Similarly, after reversal, when the desired instrumental behavior, go, now aligned with the outcome present 'win', the participants were able to perform significantly above chance with high average accuracy. Although we can posit that instrumental learning was at play within both pre and post reversal, its success highly depended on its interaction with Pavlovian biases.

Go to Avoid Losing became No Go to Avoid Losing

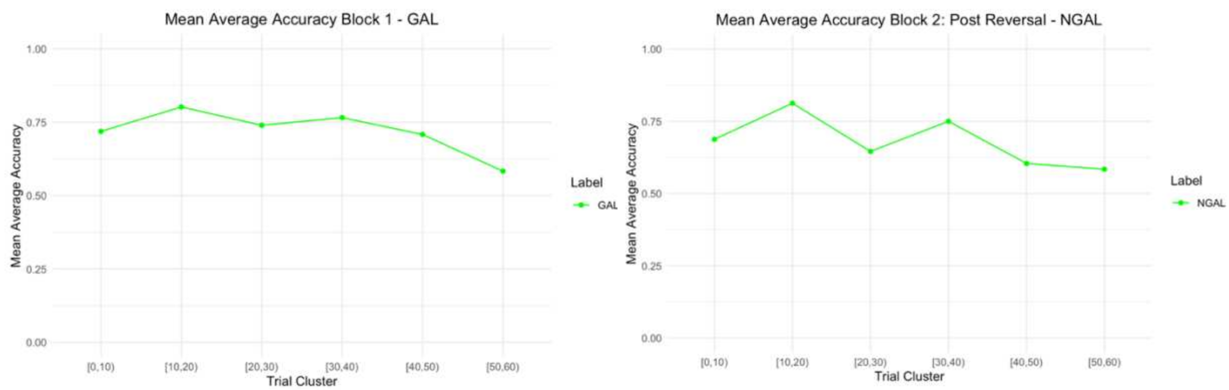
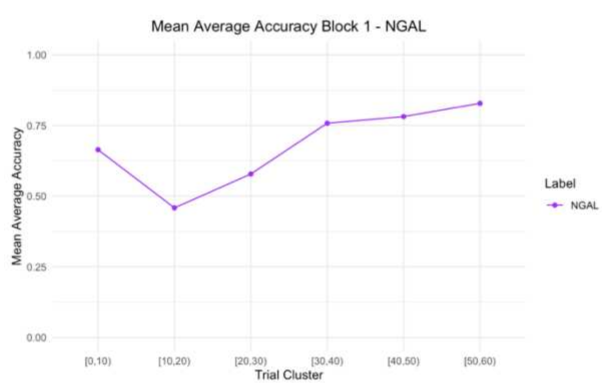


Figure 7

<u>Pre-reversal</u>: For Go to Avoid Losing (GAL) the graph indicates that, although there is a small decrease at the end, accuracy began above chance and was mostly sustained throughout the trials. This implies that although the outcome for this stimulus was not a reward, participants still performed the 'go' bias regardless of its association to the outcome – which was the correct response in this trial type. However, as they learned the Pavlovian value they became less likely to make a 'go' response and accuracy decreased. Participants' instrumental learning of action throughout this trial was impaired by the conflict with the Pavlovian bias of winning and over time participants struggled with this incongruency leading to a decline in accuracy.

Post-reversal: Post reversal GAL became NGAL where the instrumental response changed from ‘go’ to ‘no go’ for the same outcome ‘avoid losing’. We see some fluctuation within participants performance due to a switch from a more favorable bias of action over inaction however participants maintained a steady performance well above chance which denotes that they were able to learn the associations and implement them as needed. The accuracy initially was high as well, elucidating the rapid switch of instrumental learning by participants. This suggests that participants were able to learn the pre reversal Pavlovian value of the robot then were able to apply goal-directed learning to learn the new compatible response after transfer. Within the task we do see a clear reliance on the ‘go’ bias from participants than ‘no go’ evident by the unstable performance post reversal (NGAL).

No Go to Avoid Losing became Go to Avoid Losing



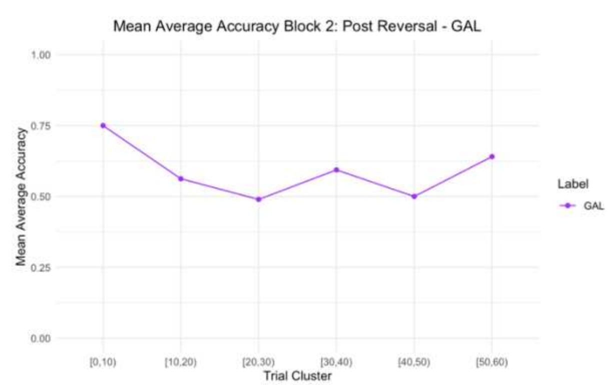


Figure 8

Pre-reversal: Participants attended to "No Go to Avoid Losing" (NGAL) stimuli during the pre-reversal phase. The graph demonstrates that although participants' initial accuracy on the first 10 trials was rather high, there was an immediate decline in accuracy in the early trials to chance levels. This implies that the idea of inaction ("No Go") in order to prevent an undesirable consequence (losing) was initially difficult for the participants to learn. However, with time participants were able to associate the two, evident with the steady increase in the accuracy pre-reversal.

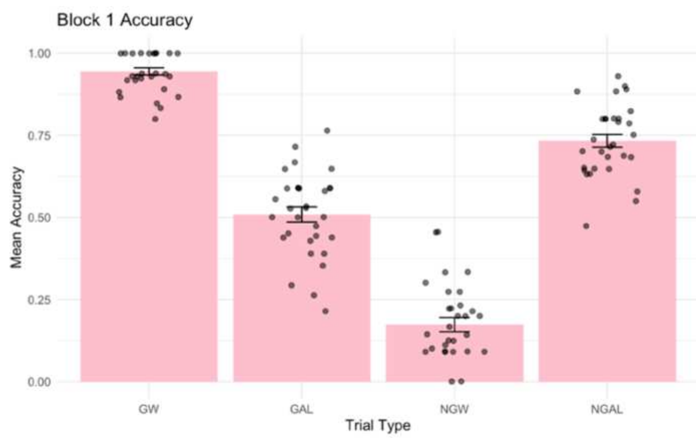
Post-reversal: Following the reversal, "Go to Avoid Losing" (GAL) became the behavior associated with the robot. Accuracy was high at first, suggesting that participants could adjust to a task where they had to take action ("Go") to prevent an aversive outcome more easily. The accuracy in the middle trials, however, shows a noticeable fall, which may suggest that individuals were having some trouble sustaining this new link over time. This could be due to the overpowering 'avoid loss' bias that was hindering the instrumental condition of 'go'. Towards the end, there is some recovery in the accuracy, indicating that individuals may have readapted and relearned their instrumental behavior. However, the initial decrease in accuracy is indicative of participants encountering challenge to reverse the contingencies Their inability to quickly switch the action needed for the new condition suggests the overpowering biases that played a key role in sustaining rigidity and inflexibility which participants were not able to override through goal-directed actions.

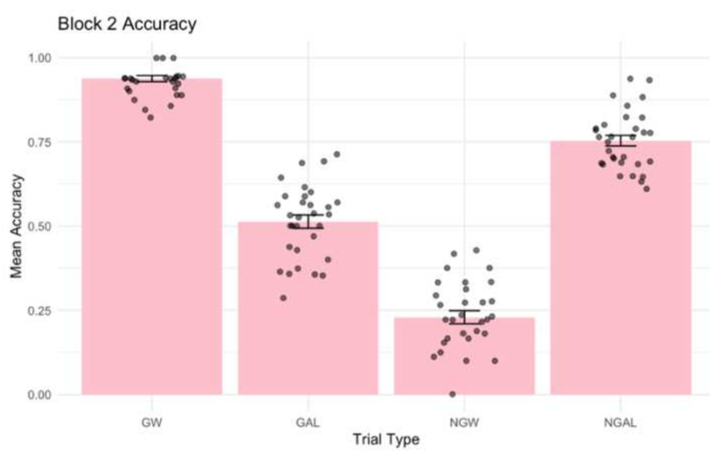
Discussion of Results from Reversal of Instrumental Learning

These results illustrate how the interaction between Pavlovian biases and goal-directed instrumental learning affect a participant's ability to adapt and switch to altered conditions. Although we saw a strong effect from the Pavlovian biases in all four reversals, we also saw suggestions of flexible goal-directed instrumental learning, especially in the reversal from GAL to NGAL. We noticed a pattern of behaviors where participants faced with congruent conditions (GW & NGAL) tend to learn the reversals of the instrumental conditioning faster as the bias attached to the action put the instrumental learning into a compatible context for participants. For example, the switch from NGW to GW was seen to have significantly high accuracy after the reversal which could be indicative of the bias putting the action (go) into a compatible context (win) leading to higher accuracy. Next, we observed that participants showed a bias towards action (go) over inaction (no go). This could be seen in a minimal disparity between GAL and NGAL, where participants exhibited almost similar performance (Figure 7 & 8) even though in a congruent condition such as NGAL —where inaction is required to avoid a loss—they would perform better due to the typically stronger, intrinsic bias toward avoiding losses through inaction. Participants showed reliance on their declarative memory as they were able to reverse conditions rapidly and during the reversal of instrumental conditions, participants exhibited a flexible and adaptive approach, however, their ability to perform or withhold an action was heavily influenced by the positive or negative value (valence) associated with the outcome.

Experiment 1c: Reversal of Pavlovian Biases:

For Pavlovian biases, we predicted significant shifts in average accuracy as the outcomes associated with each robot were reversed (e.g., 'Go to Win' changed to 'Go to Avoid Losing' and vice versa). Mean accuracy in each condition over blocks is shown in Figure 9. The results of a 3 (Block) x 4 (Trial Type) two-way ANOVA showed significant main effects of block ($F(2,5568) = 4.47$, $p < 0.05$) and trial type ($F(3,5568) = 171.78$, $p < 0.001$) where the interaction was also significant ($F(6,5568) = 260.25$, $p < 0.001$). Specifically, the changes in task blocks reflect the participants' adaptation process to the reversed outcomes as they learned the new associations. To learn more about the substantial impacts of block on accuracy, the Scheffé post hoc test was conducted. Our findings revealed that there was a significant difference ($p < .05$) between blocks 1 and 2. This aligns with our theory of reversal learning impacting accuracy as we introduced reversal within block 2. Additionally, for trial types, we found significant differences between GAL and NGAL ($p < .001$) and NGW and NGAL ($p < .001$) where NGAL reported higher average accuracy throughout the task. Additionally, we saw a significant difference between GW and NGAL ($p < .05$) as well as GW and NGW ($p < .001$) with GW's accuracy higher in comparison to both trial types. These results demonstrate the significant effect the reversal had on participants performance and the potential challenges they exhibited depending on congruency of contingencies (GW, NGAL) and incongruency as well (NGW, GAL).





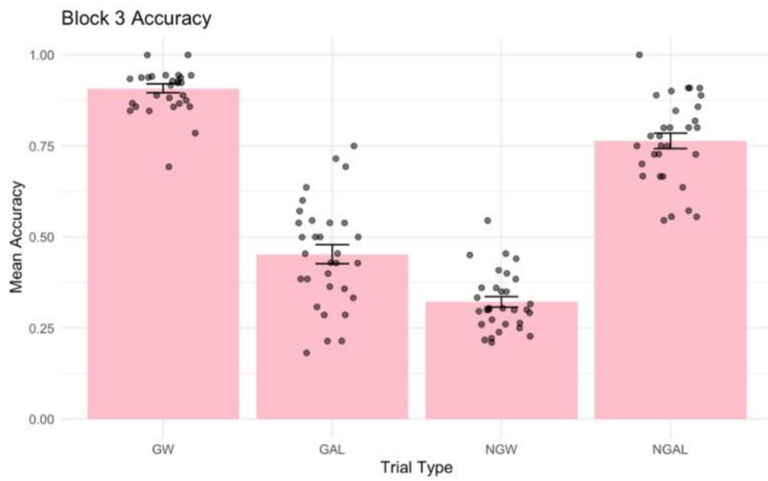


Figure 9: Experiment 1c/ Reversal of Pavlovian Bias. Average accuracy reported of participants in each block with individual data points. Valence was switched in block 3 from 'win' to 'avoid loss' and vice versa.

Focused Examination of Each Reversal Type

For the reversal of Pavlovian biases, we compared the results of each stimulus or robot pre and post reversal across mini blocks of 10 trials. We analyzed how participants performance changed before reversal set in and after the participants encountered a change of conditions which allowed us to evaluate if participants were able to adapt to the reversal or if their biases countered against their ability to adapt to reversal.

Go to Win changed to Go to Avoid Losing

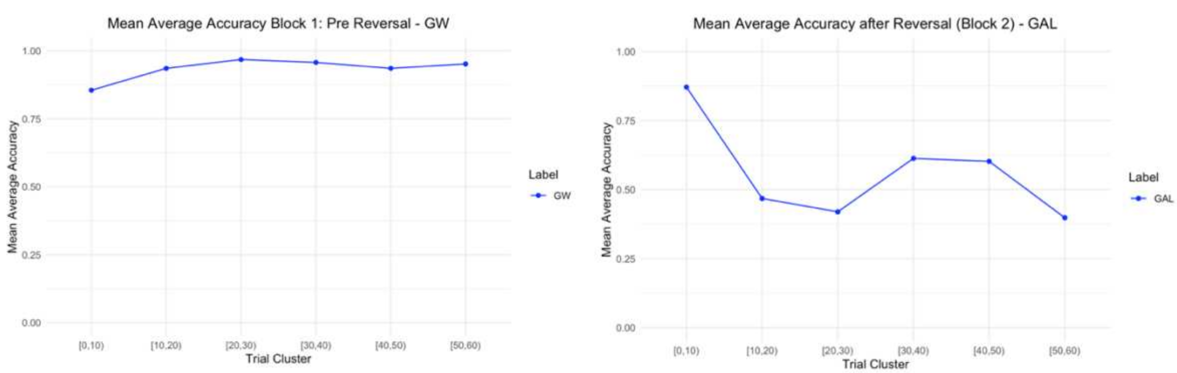


Figure 10

Pre reversal: Go to Win (GW) accuracy was high throughout the block, consistent with the congruent interaction between instrumental learning and Pavlovian bias as winning by implementing action. Participants in the pre-reversal phase showed a high mean average accuracy and consistently performed well in the 'Go to Win' (GW) condition (Figure 9 & Figure 10). This consistent performance can be attributed to the ‘Go’ action's alignment with instinctive Pavlovian biases, in which acting in order to receive a reward is a predisposition.

Post reversal: After reversal, the Pavlovian outcome was reversed (win to loss) but the instrumental response remained the same (Go). Immediately after the reversal, accuracy was high indicating that participants continued to make the go response to the robot. However, accuracy soon decreased to chance levels. This pattern suggests that the Pavlovian value had an impact on the participants performance because if they only learned the instrumental value, they would have continued with the go response which would have resulted in high accuracy throughout the block. This finding also indicates that participants were able to quickly reverse the Pavlovian value because if they had not reversed the value, they would have kept treating the robot as a "win" and accuracy would have remained high as well. This could point to some goal-directedness within the biases of participants in this task as the Pavlovian biases were malleable and could be quickly redirected. This then suggests two possibilities regarding the high accuracy within GW pre-reversal. Firstly, participants were motivated by the Pavlovian cue (win) which led them to press the button ‘go’ which underscores the absence of the instrumental learning. Secondly, perhaps, instrumental learning ran congruently with the Pavlovian biases causing participants to reach almost perfect accuracy.

Go to Avoid Losing (GAL) became Go to Win (GW)

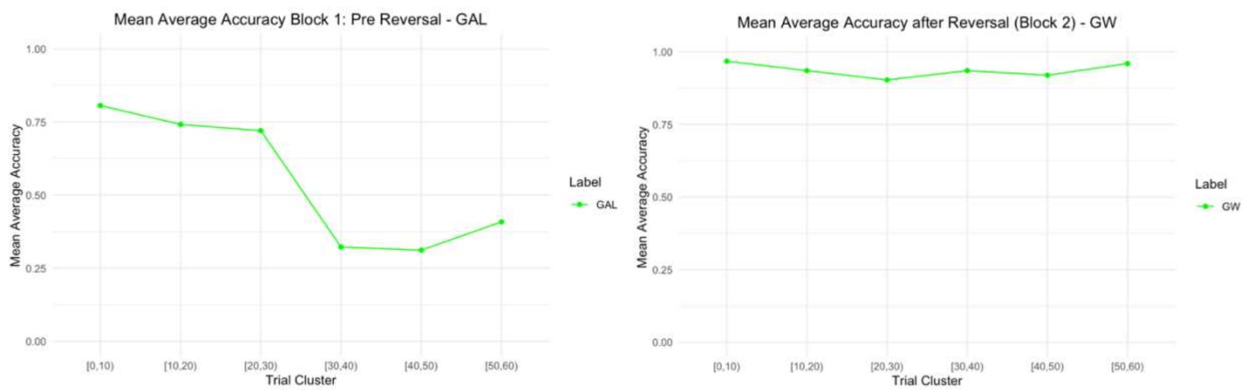


Figure 11

<u>Pre-reversal</u>: The instrumental "Go" response stayed constant pre and post reversal, but the Pavlovian association changed from an inconsistent "Loss" to a consistent "Win”. Within the pre reversal condition, participants became worse over time. Their Pavlovian biases of ‘avoid losing’ impacted their instrumental response of ‘go’. In the beginning of the task, participants performed more accurately but in later trial clusters their performance was compromised and ultimately was below chance (50% accuracy). This pattern is consistent with the general ‘go’ bias observed in this task, as well as with participants learning the robot's Pavlovian value and exhibiting a greater Pavlovian effect over trials. Given that go and loss are incompatible in this condition, the Pavlovian effect manifests itself as a decrease in accuracy.

<u>Post-reversal:</u> Post reversal, GAL became GW or Go to Win. Figure 11 shows how accuracy quickly escalated to almost perfect post reversal. This could be due to a couple of reasons. One reason could be that participants already had learned the ‘go’ response as a part of their instrumental learning but weren’t able to express it until the reversal occurred. The reversal might have made a context shift within the participants that elicited their instrumental response along with a newly shifted Pavlovian reversal to ‘win’ which evidently supported their responses

to sustain a higher accuracy. Another interpretation is that Pavlovian bias was the main factor influencing the participants' performance. Following the reversal, they promptly changed the valence from "loss" to "win" in order to conform to the new Pavlovian alteration. Because participants rapidly matched their behavior with the new, constant "win" value, the instant boost in accuracy would be explained by this rapid adaption of the Pavlovian bias. Finally, we theorize that potentially participants reverted back to the overall ‘go’ bias when the reversal occurred meaning participants were more inclined to take action over inaction post reversal as it aligned with ‘win’ bias.

No Go to Win became No Go to Avoid Losing

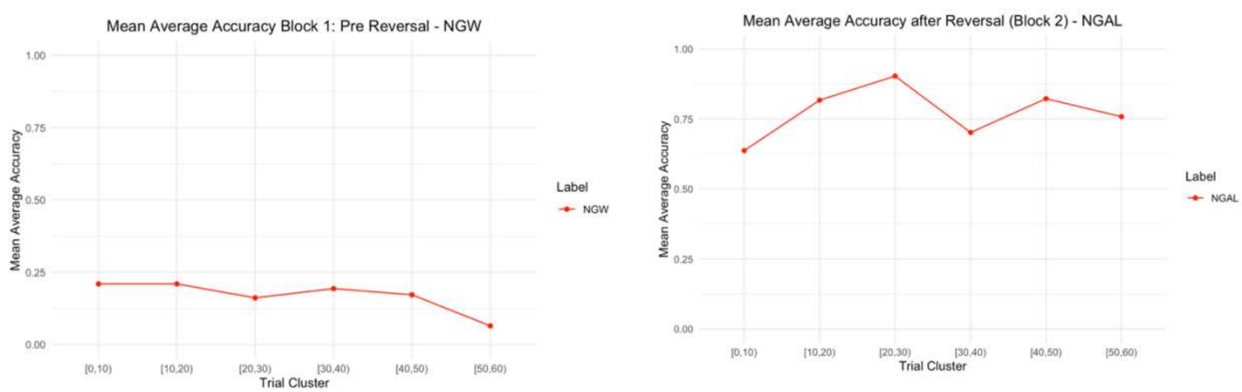


Figure 12

<u>Pre-reversal</u>: Figure 12 shows pre-reversal accuracy in NGW was below chance from the beginning of the block and decreased over training. Accuracy below chance indicates that participants were more reliant on the Pavlovian bias of ‘win’ than the instrumental action ‘No Go’. The line graph shows the inability of participants to learn the instrumental response: if they just guessed then their accuracy would have been around chance or 50% however their accuracy was significantly lower. Due to the discrepancy between the "no go" reaction and the "win" value in the pre-reversal NGW condition, accuracy decreased as the Pavlovian effect grew stronger.

Post-reversal: After reversal NGW changed to NGAL, which aligned both the instrumental learning 'no go' and Pavlovian bias 'avoid losing' which led them to avoid action to avoid a loss. Therefore, participants' accuracy significantly increased and was above chance which is indicative of their learning the instrumental response of no go but also is attributable to the switch in the Pavlovian bias.

No Go to Avoid Losing changed to No Go to Win

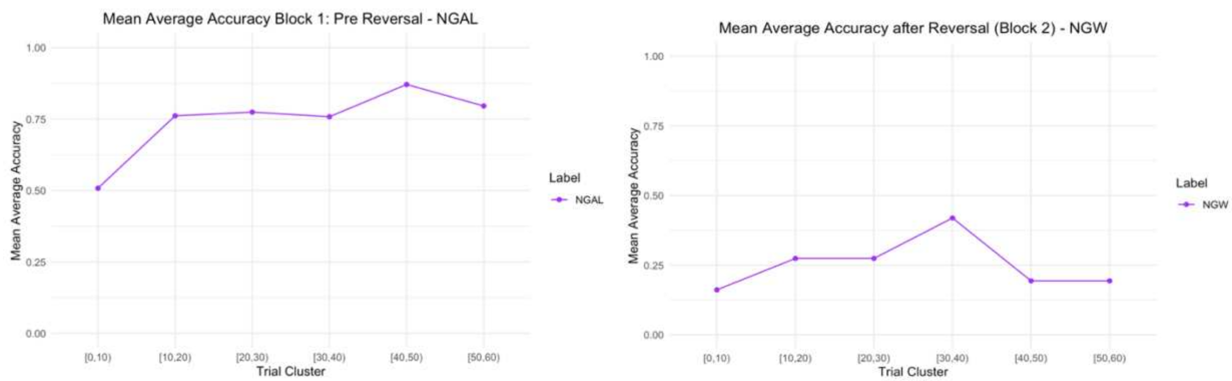


Figure 13

Pre-reversal: The "No-Go to Avoid Losing" (NGAL) condition demonstrated reasonably constant accuracy during the pre-reversal phase. Suppressing action in order to prevent a loss is in alignment with preexisting biases and consistent with the finding that participants' accuracy gradually increased as the block continued. Accuracy was far above chance even though it was not as strong as in another consistent pairing (Go to Win), indicating a potential overall bias for go responses over no-go responses.

Post-reversal: Following the reversal, the trial type shifted from NGAL to NGW. To attain high accuracy participants had to withhold response to be rewarded. Accuracy was below chance even at the beginning of the post reversal block. This underscores that individuals were affected by the Pavlovian valence of 'win' which had now become inconsistent with the action, 'no go'. The anticipation of reward in this task invigorates the biases leading participants to act, even though inaction would be the optimal option.

Discussion on Reversal of Pavlovian Biases

These results signify a strong Pavlovian influence on participants' performance. Pavlovian biases were especially influential when there was incongruency between the instrumental response required and the outcome associated with ‘win’ or ‘avoid losing’. Participants showed a strong tendency to ‘go’ with a ‘win’ bias and withhold their response during a ‘avoid loss’ bias. Additionally, the biases were more pronounced in the ‘go’ than in the ‘no-go’ conditions, indicating that the Pavlovian connections may have been amplified by an inclination toward action over inaction. Pavlovian biases appeared to have the greatest effect on reversal when the Pavlovian values (e.g., GW to GAL) were switched rather than the instrumental response (e.g., GW to NGW). For example, accuracy decreased after reversal in the GW to GAL condition, indicating a discrepancy between the initial instrumental response and the updated Pavlovian connection. This reversal would not have affected accuracy if participants had utilized goal-directed instrumental learning and continued to execute the same response. Another noticeable effect was the quick ability of participants to reverse Pavlovian biases. For example, NGW changing to NGAL (Figure 12) saw a significant increase in the average accuracy, indicative of participants' flexibility and sensitivity to a change in these biases. If instrumental learning was being utilized, then pre reversal performance would have been comparable to post reversal performance. Although there were significant Pavlovian effects as we hypothesized, it is still unclear how much instrumental learning—that is, learning to execute particular actions like Go or No Go—added to performance, especially in the ‘Go to Win’ situations where accuracy was high shortly after the reversal. A general inclination toward action over inaction might possibly be the cause of this pattern.

General Discussion Experiment 1

Across the three conditions, namely: baseline (1a; replication of Raab & Hartley), reversal of goal-directed instrumental learning (1b), and reversal of Pavlovian bias (1c) we saw consistent patterns of Pavlovian bias and how they facilitate and hinder goal-oriented actions. There were three key findings that we found with Experiment 1. One of the most salient findings from this task encompassed the relationship between Pavlovian bias and instrumental learning. We found that Pavlovian biases were present within the task, often overpowering goal-directed instrumental learning. This suggests that these biases can significantly influence behavior, even in situations where instrumental learning is expected to dominate, for example ‘No Go to Avoid Losing’. Secondly, goal-directed instrumental learning occurred at a slower pace, and in some cases of reversal, it was unclear whether the observed behavior was driven by true flexible instrumental learning or simply by the inherent ‘go’ biases. Teasing out the two behaviors and their overall effect on participants performance posed a challenge. Lastly, there was a noticeable ‘go’ biases overall, as evidenced by stronger learning in the ‘go to win’ condition compared to the ‘no-go to avoid losing’ condition. This ‘go’ bias that showed significant persistence throughout the task and was able to be readily employed, more so than inaction.

Limitations and Future Directions for Experiment 1

These results reveal several intriguing patterns. However, there are also limitations that need to be noted. Firstly, participants showed low levels of instrumental learning. In instrumental learning, an action (such as "Go" or "No Go") is linked to an outcome (such as win or loss). It is doubtful that participants would be able to adjust their behavior in response to task modifications (e.g., when the relationship between an action and its consequence was altered) if they were unable to acquire the original associations in an efficient manner. Secondly, we noticed participants being influenced by the Pavlovian values or valence of each trial type – even after reversal. Perhaps this suggests that rather than procedural mechanisms, these effects may instead be the result of conscious application of declarative knowledge. Declarative knowledge refers to information that is consciously known, such as knowing that a certain stimulus is associated with a reward. The ability of participants to quickly reverse their Pavlovian biases raises the possibility that they were depending more on declarative information than on implicit habitual reactions. Another limitation is that Experiment 1 was subdivided into three main studies that were conducted independently and in sequence (1a first, then 1b, then 1c). Because participants were not randomly assigned to these studies, we can't be certain whether differences between the studies are due only to the manipulations within the studies or due to other potential confounds (e.g., differences in the participant population at different times in the semester). In Experiment 2 we randomly assigned participants to conditions to account for this limitation.

To specifically address concerns about limited learning and noise within our results from Experiment 1, we conducted an extension to Experiment 1 by deliberately altering the task to make it easier to learn for participants and extended the amount of training before reversal. With the new modifications in Experiment 2, our aim was to ensure that participants have learned all the associations (as evidenced by accuracy higher than chance) before reversal. We altered the task in such a way to engage participants to learn the conditions of each robot fully to ensure we attain robust findings that can assist us in exploring the effects of reversal on the go/no-go task.

EXPERIMENT 2

Experiment 2 incorporated three key manipulations that built upon the already established Experiment 1: baseline (2a), reversal of goal-directed instrumental learning (2b) and reversal of Pavlovian biases (2c). With Experiment 2, our primary aim was to refine our paradigm in order to optimize learning for our participants. With enhanced learning, we were able to examine the interactions between goal-directed instrumental learning and Pavlovian biases in depth and achieve more robust findings. As in Experiment 1, our main goal was to examine the flexibility Pavlovian biases exhibit when contingencies change. For this experiment we hypothesized that post reversal Pavlovian biases will not be able to adapt and impede participants performance due to their dependency on implicit and procedural learning mechanisms whereas goal-directed instrumental learning, reliant on declarative and conscious action-taking would be quick to adapt to the reversal.

Methods

Participants

After conducting a power analysis, we recruited approximately 90 participants (30 in each condition) for Experiment 2. The power analysis showed that for a significance level of 0.05, and a desired power of 0.8, the study was sufficiently powered to detect the effect size of 0.49. We recruited from the same population pool as Experiment 1 at Colorado State University, specifically from PSY100 & PSY250 courses to ensure no additional variability was introduced to the data. Participants ranged from 18 – 23 years old (M = 18.87; SD = 1.50) with 24 males, 59 females and 2 non-binary persons in the dataset.

Behavioral Task Design

For Experiment 2 we modified our paradigm utilized in Experiment 1 to achieve better accuracy (greater than chance, which was 50%) and reduce noise within our dataset. The primary task for Experiment 2 was the Pavlovian go/no-go task, inspired by the version designed by Raab & Hartley (2020) and the baseline task utilized for Experiment 1a during task reversal paradigm. In the modified version of the paradigm, we made the following alterations to improve overall accuracy and reduce noise. First, the relationship between stimuli and outcomes that was set within the baseline condition (Experiment 1a) and subsequently maintained within the reversal conditions was changed from 80 - 20% probabilistic to 100% deterministic in order to make the task easier to learn for participants. Secondly, we reduced the duration of each trial in order to accommodate an extra training block during our one-hour testing sessions for participants to better learn the associations of each robot prior to reversal. We anticipated that shortening the trial length would also engage the participants better which would result in higher accuracy levels. The reduction of trial time was implemented by shortening the time of the fixation points and outcome duration (Figure A). Thirdly, we added an extra learning block before reversal occurred, making it a total of 4 blocks with the first three blocks for initial learning, and block 4 for reversal. Lastly, we added a post-task survey (refer to Appendix) that assisted us in understanding the strategies that participants used in learning and their explicit knowledge about each robot's associations with go versus no-go responses and win versus loss outcomes. To summarize, we included four major changes to the task, namely: shift to deterministic relationships, task duration alteration, addition of an extra learning block before reversal, and addition of a post-task survey to assess participants' knowledge of the paradigm.

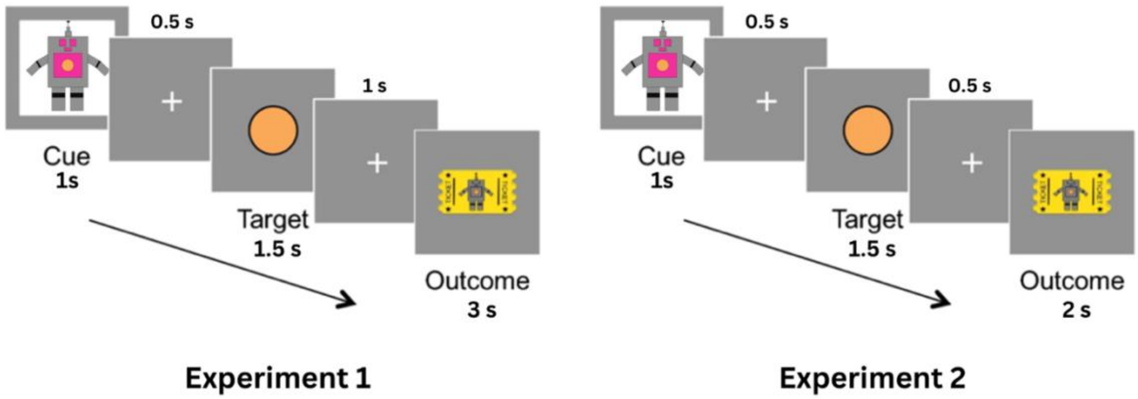


Figure A represents modifications in trial times from Experiment 1 to Experiment 2 with specific changes to fixation point and outcome time.

Results

Experiment 2a: Modified Task Results: No Reversal (Baseline) Condition

The alterations made to the original paradigm resulted in significant shifts in overall accuracy, however, the patterns of differences across task conditions remained the same. Our aim with the modifications was to achieve a higher accuracy level within the task, with mean accuracy in all conditions greater than chance (>50%). This goal was successfully achieved. As shown in Figure 14 performance in all conditions was now greater than chance levels.

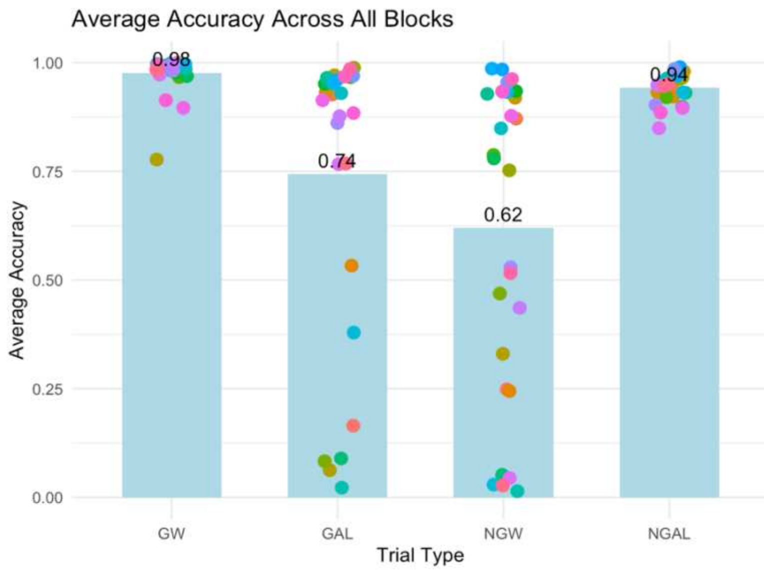


Figure 14: Average accuracy performed on each trial type across all blocks in Experiment 2a (baseline) with individual data points representing individual participants' performance.

We observed similar qualitative patterns in the results as in our original task (Experiment 1a). Contingencies that were congruent with natural tendencies—such as acting to obtain a reward (GW) and withholding a response to prevent an undesirable outcome (NGAL)— resulted in best performance from participants. Conversely, stimuli requiring behaviors that were incongruent, like acting to prevent losing (GAL) or withholding action to win (NGW), resulted in worse performance.

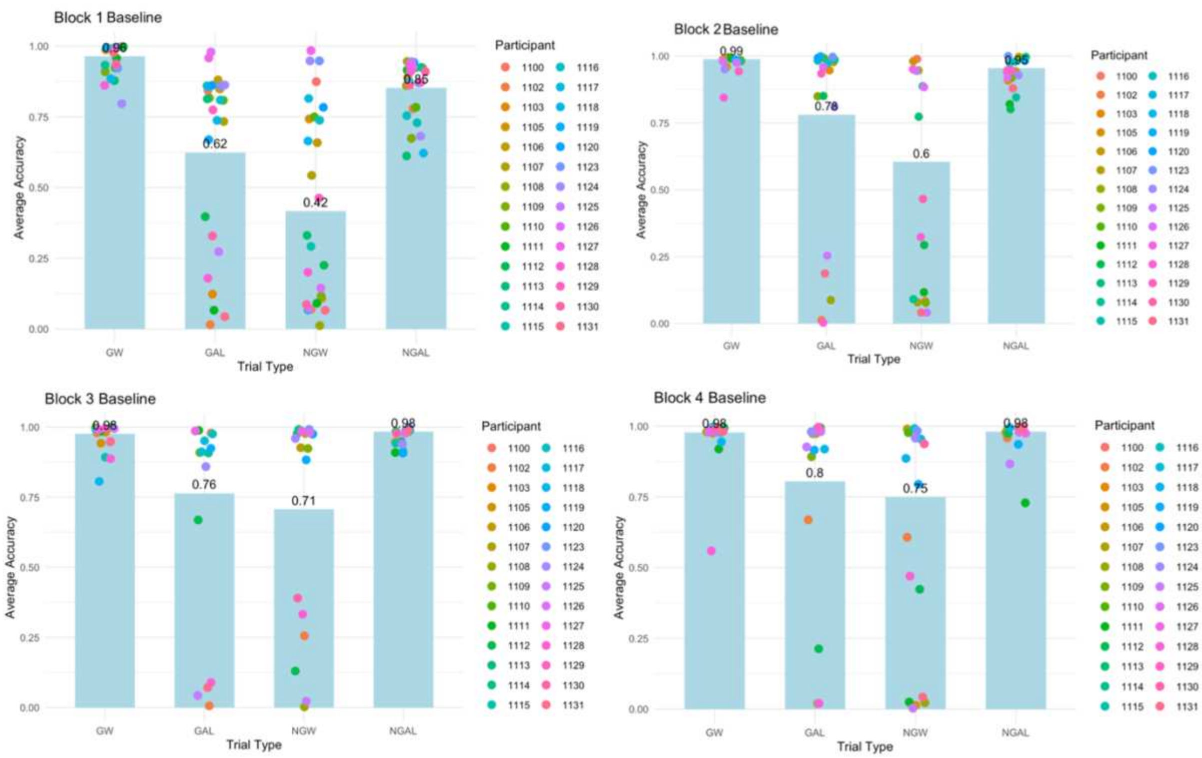


Figure 15: Experiment 2a. Accuracy in each block with individual data points from participants in modified baseline task (GW, GAL, NGW & NGAL).

Figure 15 demonstrates the mean accuracy of trial types within each block. The graph reveals significant increases in accuracy for GAL and NGW across blocks. The average accuracy for GAL and NGW were around chance with significant subset of population performing below 50% accuracy in block 1. However, by block 4 we see an increase in accuracy for the aforementioned trial types, now at 80% for GAL and 75% for NGW and almost all of the participants improving.

To analyze the effects of block and trial type (GW, GAL, NGW, NGAL) and their interaction on accuracy we conducted a 4 (trial types) x 4 (blocks) two-way ANOVA. Both block ($F(3,6677) = 513.51$, $p < .001$) and trial type ($F(3,6677) = 96.95$, $p < .001$) showed significant main effects in the analysis indicative of their significant effects on accuracy. There was also a significant interaction ($F(9, 6677) = 16.48$, $p < 0.001$) between the block and trial type. To explore these effects in-depth post hoc pairwise comparisons were conducted using the estimated marginal means (EMM) for our two-way ANOVA. Post hoc pairwise comparisons revealed significant differences between pairs in each block. For example, all trial types were significantly different from each other from block 1 through block 4 except for trial types of GW and NGAL, where they did not differ significantly in any blocks due to high average accuracy (almost 100%) throughout the task. However, trial types such as GW (block 4) and NGW (block 4) were significantly different along with NGW (block 4) and NGAL (block 4), although there were no significant differences between GW and NGAL in block 4 ($p = 0.99$).

Experiment 2b: Reversal of Goal-directed Instrumental Learning:

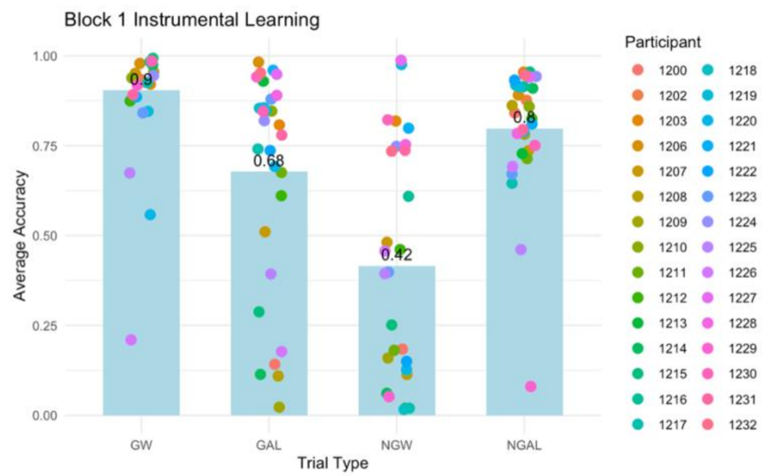
After modifying our paradigm for Experiment 2, we saw significant changes within participants' performance in comparison to Experiment 1. The reversal of goal-directed instrumental learning encompasses the same switches as Experiment 1b in which we altered the actions associated with each robot (for e.g. ‘Go to Win’ became ‘No Go to Win’). The reversal was implemented after three out of the four blocks of training. For this condition we hypothesized that participants will be able to adapt to the changing contexts for reversal of goal-directed instrumental learning better post reversal due to their reliance on declarative and flexible mechanisms (Experiment 2c).

We first examined the pre reversal blocks. Similar to the baseline condition (Experiment 2a), participants demonstrated a clear learning trend from block 1 to block 3. Additionally, we observed the pervasive Pavlovian bias trend in the pre reversal blocks. As shown in Figure 16, we see that in block 3 there is an increase in performance on all four trial types, with their average accuracy significantly above chance level (>50%) in each. This trend laid the foundation for our reversal to take place in block 4, where participants now had to change the associations of valence they attributed to a specific robot in the previous blocks. For our pre reversal analysis, we chose to conduct a two-way ANOVA which reported significant main effects of block ($F(2, 5388) = 45.84$, $p < 0.001$) – suggesting that average accuracy of participants varied between each of the three pre reversal blocks – and of trial types ($F(3, 5388) = 180.49$, $p < 0.001$) suggesting variance in performance for the four trial types (GW, GAL, NGW and NGAL). Lastly, the interaction between block and trial type also significantly affected participants performance on the task ($F(6, 5388) = 10.97$, $p < 0.001$). Additionally, after conducting post hoc pairwise test, we found that GW was significantly different than NGW and GAL ($p < .001$), with superior performance throughout the pre reversal blocks (block 1, block 2 and block 3). Additionally, GW (higher accuracy) and NGAL which were significantly different ($p < .001$) in block 1, block 2, and block 3. Overarchingly, the analysis reported significant differences in most trial type pairs with NGW with significantly lower accuracy.

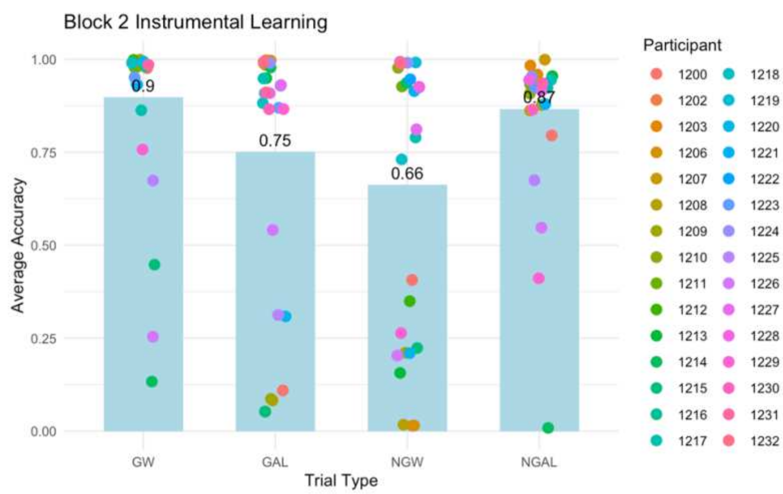
To account for block 4 (post reversal) we conducted an additional one-way ANOVA which also revealed significant effects of trial types ($F(3, 1796) = 29.16$, $p < .001$) on accuracy of participants. We then conducted a post-hoc through estimated marginal means (EMM) test to understand the interactions between trial types and blocks through pairwise comparisons adjusted through Tukey HSD method to control for family-wise errors. The post-hoc test revealed significant differences amongst the four trial types through the block. As can be seen in Figure 16, the most theoretically important differences in block 4 (post reversal) were between GAL and GW ($p < .001$) with GW acquiring higher accuracy. Additionally, GAL and NGAL also differed significantly ($p < .001$) as GAL had a significantly higher accuracy in block 4.

However, GW and NGAL did not differ significantly ($p = .864$) as well as no significant differences were reported between trial types GAL and NGW ($p = .123$).

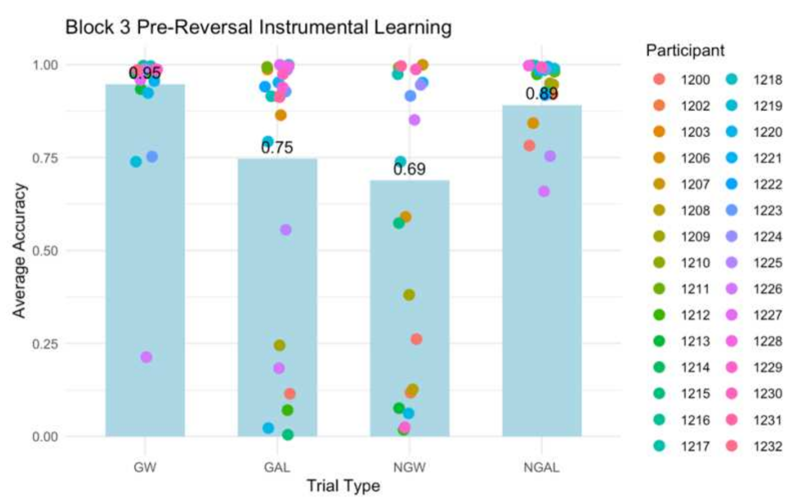
We supplemented our ANOVA results with linear mixed effects model (LMM) to test how our fixed effects, trial and blocks, and our random effects of participant variability affected performance. In this analysis, LMM appoints one of the levels of the variable as a baseline which is necessary for interpretation and comparisons. The results indicated significant main effects of block and trial type (GW, GAL, NGW, NGAL). Specifically, performance decreased significantly in post reversal in block 4 compared to the baseline (block1) ($p < .001$), while block 3 ($p = .073$) and block 2 ($p = .777$) showed no significant differences from block 1 which was the putative baseline for this analysis. Within trial type, NGW, NGAL and GAL performed significantly lower compared to GW (all $p < 0.001$). The ability for LMM to account for individual differences assisted us to account for random intercepts as well. This analysis revealed participant specific variability of 0.031 ($\sigma = 0.177$). This analysis indicates that participants varied 0.031 in their baselines when compared to each other while standard deviation of magnitude 0.177 which is relatively small compared to our entire model. This random effect value considers participants performance individually and which participants are able to perform superiorly or inferiorly from others – this was accounted for to increase the accuracy of fixed effects.



Block 1 Instrumental Learning
Average Accuracy
1.00
0.75
0.50
0.25
0.00
0.9
0.68
0.42
0.8
GW
GAL
NGW
NGAL
Trial Type
Participant
1200 1202 1203 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217
1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232



Block 2 Instrumental Learning
Average Accuracy
1.00
0.75
0.50
0.25
0.00
0.9
0.75
0.66
0.87
GW
GAL
NGW
NGAL
Trial Type
Participant
1200 1202 1203 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217
1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232



Block 3 Pre-Reversal Instrumental Learning
Average Accuracy
1.00
0.75
0.50
0.25
0.00
0.95
0.75
0.69
0.89
GW
GAL
NGW
NGAL
Trial Type
Participant
1200 1202 1203 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217
1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232

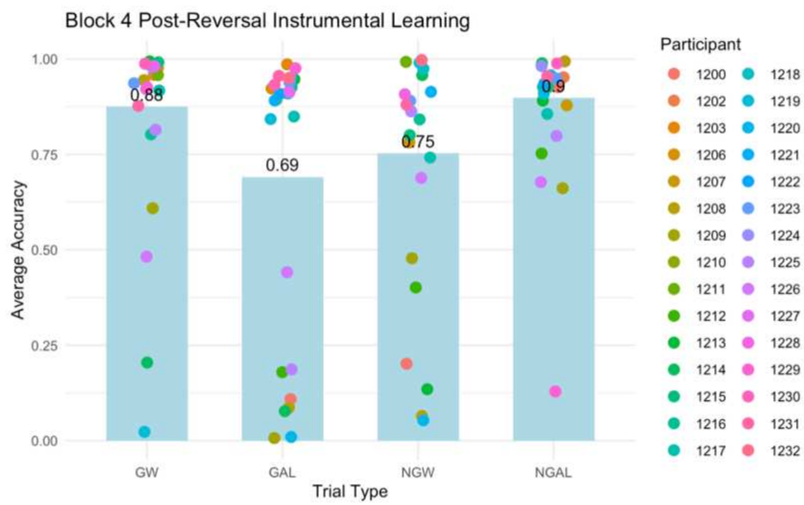


Figure 16: Experiment 2b. Accuracy in each block with individual data points from participants in reversal instrumental learning condition (GW, GAL, NGW & NGAL).

Focused Examination of Each Reversal Type

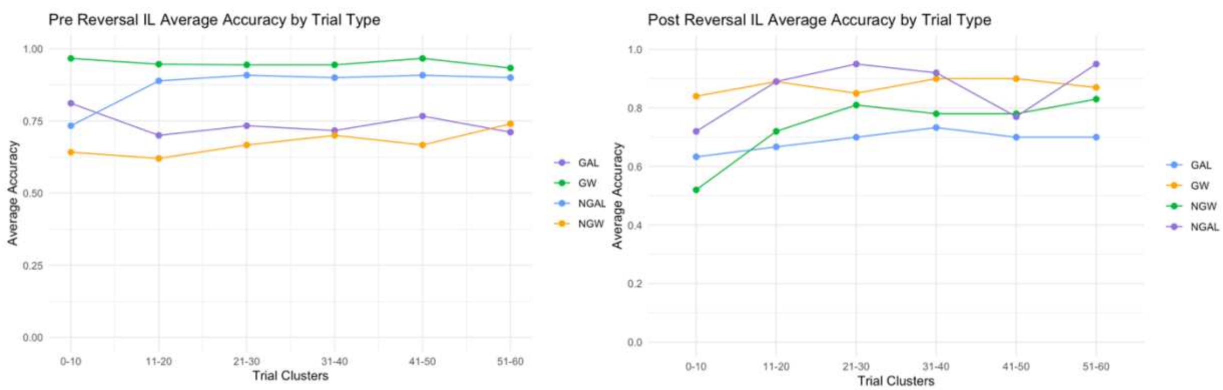


Figure 17: Trial by trial examination of Block 3 (pre-reversal) and Block 4 (post-reversal)

For the reversal of goal-directed instrumental learning, we compared the results of each stimulus or robot pre and post reversal across mini blocks of 10 trials. We analyzed how performance changed before reversal set in and after the participants encountered a change of conditions, this allowed us to evaluate if participants were able to adapt to the reversal or if their biases countered against their ability to adapt to reversal

Go to Win changed to No Go to Win

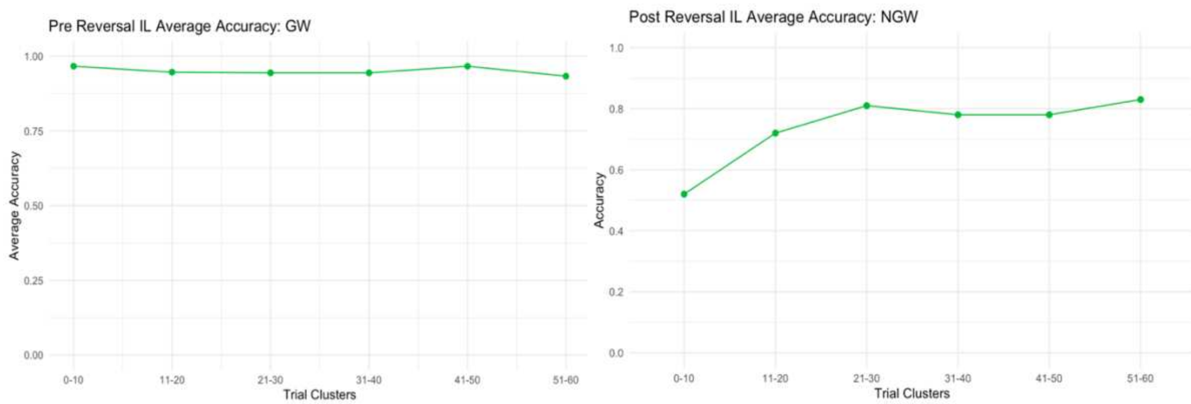


Figure 18

<u>Pre-reversal</u>: 'Go to Win' showed high average accuracy throughout the block as participants were able to sustain their performance. As mentioned before, this most likely occurred due to the alignment of both the instrumental learning response of 'go' and the Pavlovian bias of 'win'. <u>Post-reversal</u>:, 'Go to Win' became 'No Go to Win' where the instrumental learning aspect of the stimuli was now switched from 'go' to 'no-go' while the bias of 'win' remained the same. The average accuracy of participants fell around chance at the beginning of the block suggesting that their ability to associate the no-go response with a reward was compromised. However, as the trials proceeded, we saw an increase in the average accuracy showing that participants were now able to learn and perform the reversed contingencies. In line with our hypothesis, this line trend post reversal showcases the flexible nature of instrumental conditions, and the accessibility of cognitive flexibility when instrumental goal-directed behaviors need modification. However, even after accuracy plateaued, it was still lower (around 80%) than block 3 pre reversal accuracy

(near 100%) accuracy due to the lack of practice participants received for the reversed conditions (Figure 18).

No Go to Win changed to Go to Win

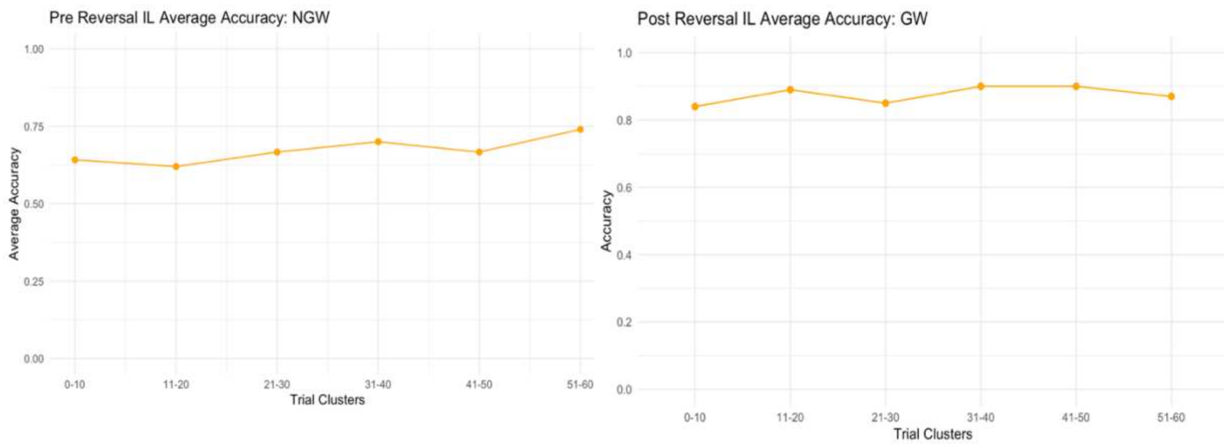


Figure 19

<u>Pre-reversal</u>: No Go to Win (NGW) maintained a low accuracy performance albeit still above chance, underscoring participants difficulty combining the incongruent instrumental learning 'no go' response and Pavlovian bias 'win' together (Figure 19). Even at the end of the block 3, after 180 trials, average performance was still only modestly above chance (60%) and a subgroup of participants were unable to learn the association, indicating an inability to associate inaction with reward (refer to figure 16 with individual data points).

<u>Post-reversal</u>: The reversal transitioned the NGW stimuli to 'Go to Win'. Figure 19 shows high accuracy immediately after the instrumental response switched from an incongruent action to a congruent one. Participants' ability to reverse the instrumental condition rapidly underscores their ability to learn the Pavlovian value of the robot, in the pre-reversal condition. Again, due to the flexibility of instrumental learning, participants were able to switch their learned responses well. This transition underscores the importance of valence and its interaction with the

instrumental response. Participants were able to learn the NGW association in the pre reversal condition, however with the alignment of the Pavlovian valence with the more compatible instrumental response (Go), we saw a significant increase in performance.

Go to Avoid Losing changed to No Go to Avoid Losing

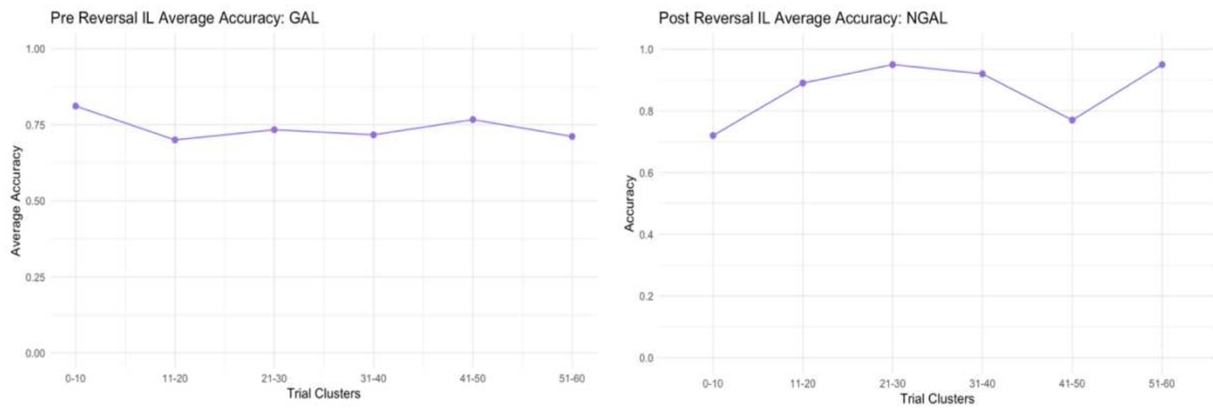
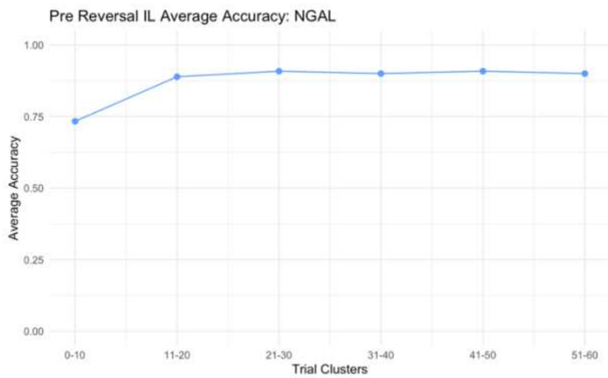


Figure 20

Pre-reversal: Accuracy in Go to Avoid Losing (GAL) was sustained at around 75% throughout block 3 which was above chance but below optimal performance. This lower accuracy is likely due to the incongruency between Pavlovian valence and instrumental action in this condition. It should be noted that in the NGW condition there was substantial between-subject variability. As shown in Figure 16, around half of the participants achieved high levels of performance, a quarter responded at around chance, and the rest of responded based solely on Pavlovian value, leading to very low accuracy in this condition.

Post-reversal: Post reversal, GAL became NGAL. After the switch the average accuracy increased indicating participants ability to modify their actions quickly. We see some fluctuation within participants performance specifically on the 41-50 trial cluster (Figure 20) but overall participants maintained a steady performance well above chance.

No Go to Avoid Losing changed to Go to Avoid Losing



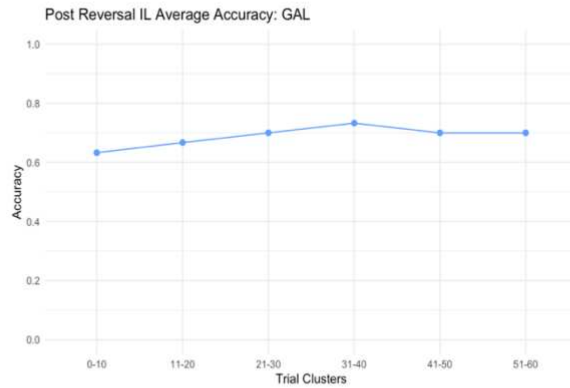


Figure 21

Pre-reversal: 'No Go to Avoid Losing' (NGAL) accuracy can be characterized by its stability and high accuracy throughout. The graph demonstrates that although participants' initial accuracy on the first 10 trials was only 75%, by block 2 performance had plateaued at around 92%. Again, we see the trend of high accuracy when instrumental and Pavlovian responses are compatible.

Post-reversal: After the reversal inaction changed to action, from NGAL to 'Go to Avoid Losing' (GAL). Overall, performance post reversal was lower in accuracy, around 70%, typical of conditions with incongruency between valence (avoid loss) and action (go).

Discussion of Results from Reversal of Instrumental Learning

The results for Experiment 2b are in congruency with our hypothesis and highlight the complex interaction between decision-making systems. With enough training on learning how to associate actions with valence participants performed significantly better on the pre reversal conditions, satisfying the changes successfully required for our Experiment 2. In all four reversals we saw the flexibility of goal-directed instrumental learning especially when the trial types transformed from incongruent conditions to congruent ones (GAL to NGAL and NGW to GW). We noticed a pattern that was exhibited in Experiment 1 as well where participants who were faced with congruent conditions (GW & NGAL) tended to learn the reversals of the instrumental conditioning faster and adapt quicker as the bias attached to the action put the instrumental learning into a compatible context for participants. Specifically, the switch from NGW to GW showed significantly high accuracy after the reversal highlighting the bias of putting the action (go) into a compatible context (win) leading to higher accuracy. Additionally, we observed that participants showed a bias towards action (Go) over inaction (No Go) which could be due to an innate desire of control and proactivity. Due to the rapid adjustment in the performance post reversal, we can infer that participant relied heavily on declarative and consciously recalled memories to reverse the conditions quickly. This can be seen throughout all four reversal conditions where participants employed an adaptive and flexible approach to quickly reverse their actions to obtain the same valence.

Experiment 2c: Reversal of Pavlovian bias

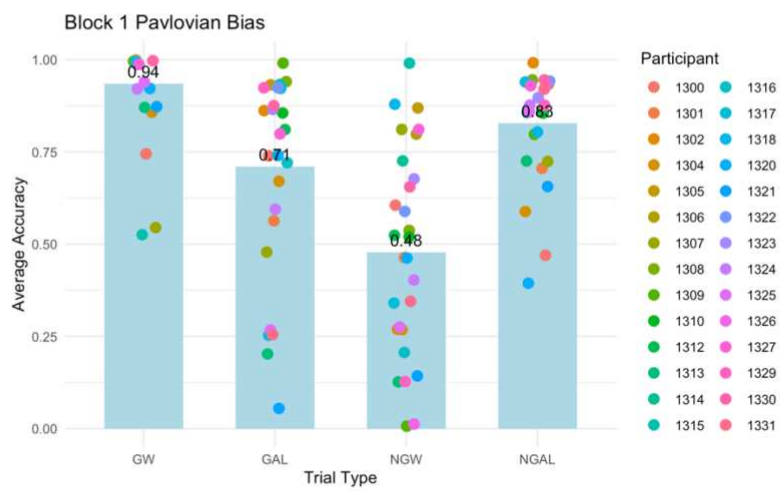
With the reversal of Pavlovian biases, the outcomes were switched (for e.g. 'Go to Avoid Losing' became 'Go to Win') in block 4. We predicted that once participants learned each Pavlovian association related to a specific robot, they would find it challenging to adapt their behavior following a reversal, resulting in poorer performance and worse transfer compared to quick adjustments made by participants in goal-directed instrumental learning.

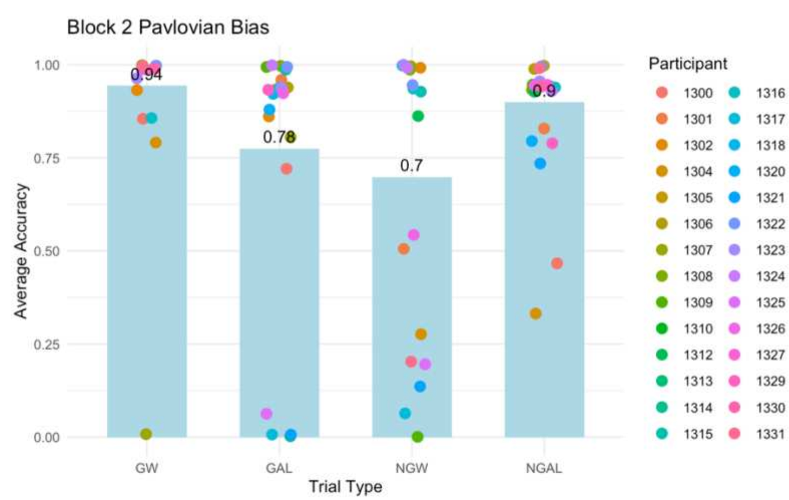
For the pre reversal, we see a clear learning trend as participants continued to learn the associations from block 1 till block 3 (Figure 22). After much practice, in block 3 we see an increase in performance on all four trial types, with average accuracy significantly above chance level (>50%). This pattern laid the foundation for our reversal to take place in block 4, where participants now had to change the associations of valence they attributed to a specific robot in the previous blocks. Mean accuracy in each condition over blocks is shown in Figure 22. The

ANOVA conducted on the pre reversal blocks (block 1, block 2 and block 3) revealed a significant main effect of trial type ($F(3,5388) = 182.98$, $p < .001$), indicating that accuracy varied significantly across trial types (GAL, GW, NGAL, and NGW). The effect of block on accuracy was also significant ($F(2,5388) = 39.03$, $p < .001$) as well as the interaction between block and trial type indicated that effect of trials on performance differed with each block ($F(6,5388) = 12.15$, $p < .001$). After conducting a post hoc pairwise comparisons through estimated marginal means (EMM), we found significant differences between GW and GAL in block 1, 2 and 3 with GW consistently attaining higher accuracy ($p < 0.001$) however found no significant differences between the two in block 3 ($p = 1.0$). We also saw GAL with superior performance than NGW ($p < 0.001$) in block 3.

Accounting for our post reversal block as well we conducted a one-way ANOVA which showed significant main effects of trial type ($F(3, 1796) = 43.42$, $p < .001$) in block 4 (post reversal block) on participants performance. We also conducted a post-hoc estimated marginal means (EMM) test to examine the pairwise comparisons extensively. The post-hoc reported significant differences between GAL and GW ($p < .001$) in addition to GAL and NGAL with GW and NGAL significantly higher respectively. There were no significant differences between GAL and NGW ($p = .727$) as well as GW and NGAL ($p = .193$). These interactions elucidated the variance in performance participants exhibited.

In addition to our ANOVA, we conducted a Linear Mixed Model (LMM) to include random effects in order to solidify our findings from ANOVA analysis. In this analysis, LMM appoints one of the levels of the variable as a baseline which is necessary for interpretation and comparisons. The random effects analysis reported a variance of 0.023 ($\sigma = 0.152$), indicative of the variability in baseline performance across the different participants. This analysis indicates that participants varied 0.023 in their baselines when compared to each other while standard deviation of magnitude 0.152 which is miniscule compared to our model. In this LMM analysis the baseline or reference level was block 1 which was significantly different from the post reversal block or block 4 ($p < 0.001$), however, predictably, block 2 ($p = .773$) did not show significant differences from block 1. Lastly, all trial types varied significantly from each other when compared to baseline trial type (GW).





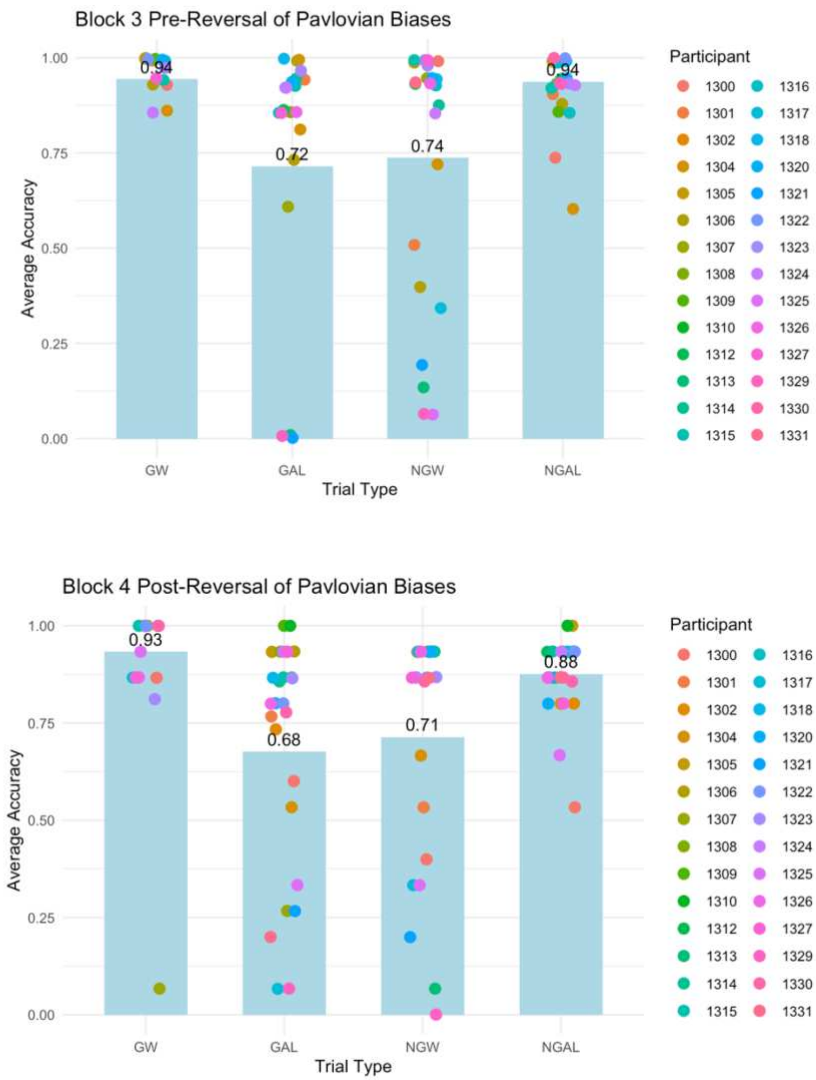


Figure 22: Experiment 2c. After implementing changes within the paradigm, a higher overall average accuracy that is greater than chance was achieved within each of the four trial types (GW, GAL, NGW & NGAL).

Focused Examination of **Each Reversal Type**

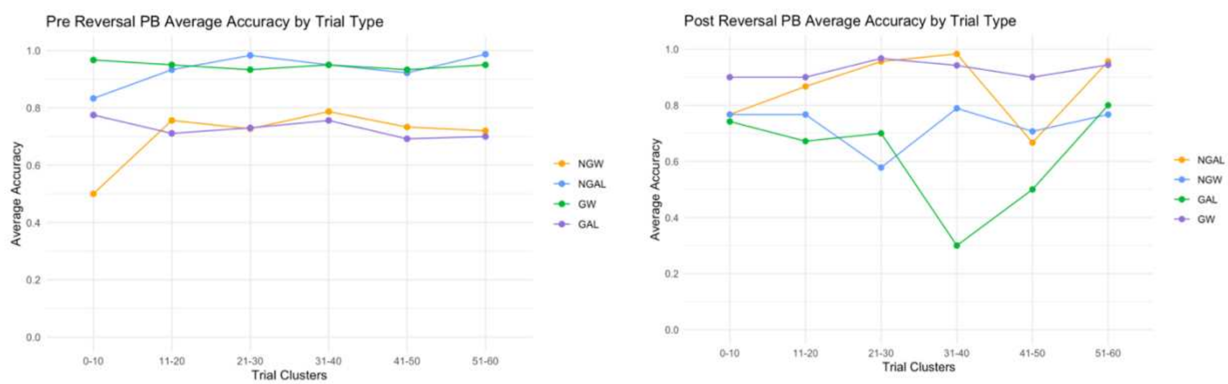


Figure 23: Trial by Trial examination of pre reversal (block 3) and post reversal (block 4)

For the reversal of Pavlovian bias, we compared the results of each stimulus or robot pre and post reversal across mini blocks of 10 trials. We analyzed how performance changed before reversal set in and after the participants encountered a change of conditions, this allowed us to evaluate if participants were able to adapt to the reversal or if their biases countered against their ability to adapt to reversal

Go to Win to Go to Avoid Losing

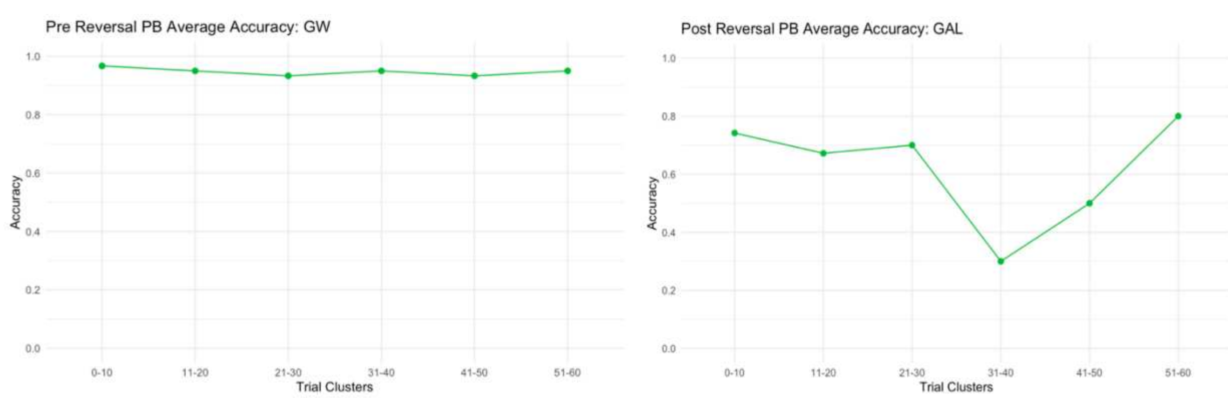


Figure 24

Pre reversal: The average accuracy of 'Go to Win' was consistently high pre reversal, as expected with the alignment of goal-directed instrumental response and Pavlovian valence. The accuracy is indicative of how well this condition aligned with the inherent bias towards reward and its association to high vigor.

Post reversal: After reversal 'Go to Win' became 'Go to Avoid Loss', where the Pavlovian bias was reversed (win to loss) but the instrumental response remained the same (go). Immediately after the reversal the average accuracy dropped but remained above chance at around 70% but accuracy continued to drop in trial cluster 31-40 but increased again in trial cluster 41-50. This pattern indicates two things, one being that participants were actually able to learn the Pavlovian value or valence modification post reversal which interfered with the 'go' action. Secondly, there is an overpowering effect of valence over instrumental response which we see by the decline in accuracy post reversal in the middle of block 4.

Go to Avoid Loss to Go to Win

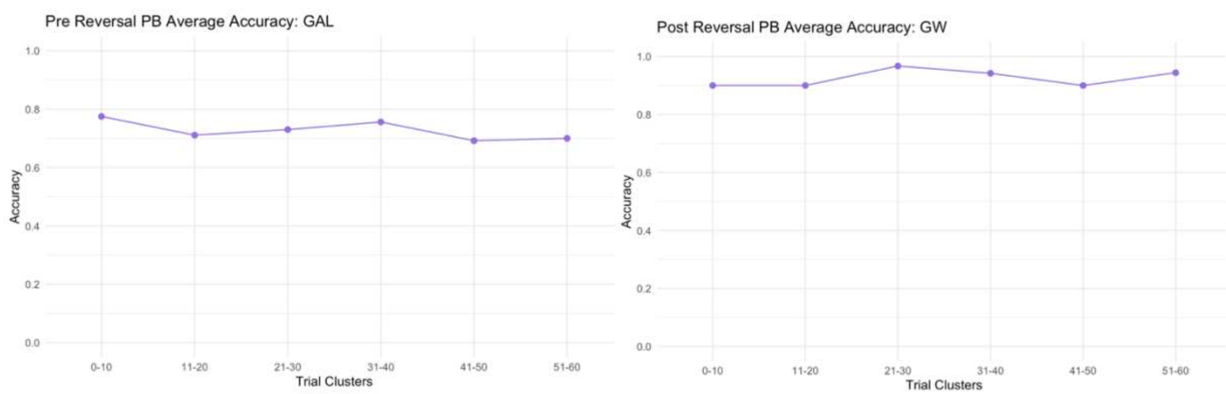


Figure 25

Pre-reversal: For the reversal of 'Go to Avoid Loss' the instrumental "Go" response was unchanged pre and post reversal, but the Pavlovian association changed from 'loss' to 'win'. Within the pre reversal condition, we saw a steady performance around 70% average accuracy with almost no fluctuations in the data, well above chance but below perfect accuracy. This accuracy is indicative of participants ability to learn associations that might not be completely congruent but are assisting them in avoiding loss.

Post-reversal: Post reversal, 'Go to Avoid Loss' became 'Go to Win'. Figure 24 shows rapid adaptation which led to a quick increase in average accuracy to almost 100% post reversal. We posit that this transition could be due to a couple of reasons. One reason could be that participants already had learned the 'go' response as a part of their instrumental learning but weren't able to express it until the reversal occurred which paired their invigoration with the reward. The reversal caused a contextual shift within the participants that elicited their instrumental response along with a newly shifted Pavlovian reversal to 'win' which evidently supported their responses to sustain a higher accuracy, even higher than pre reversal. Another reason could be that Pavlovian bias was the main factor influencing the participants' performance which led participants to quickly change their responses following reversal. The immediate increase in the post accuracy trials suggests that participants were able to use the compatibility between the valence and the go response to correctly perform the instrumental response needed.

No Go to Win to No Go to Avoid Loss

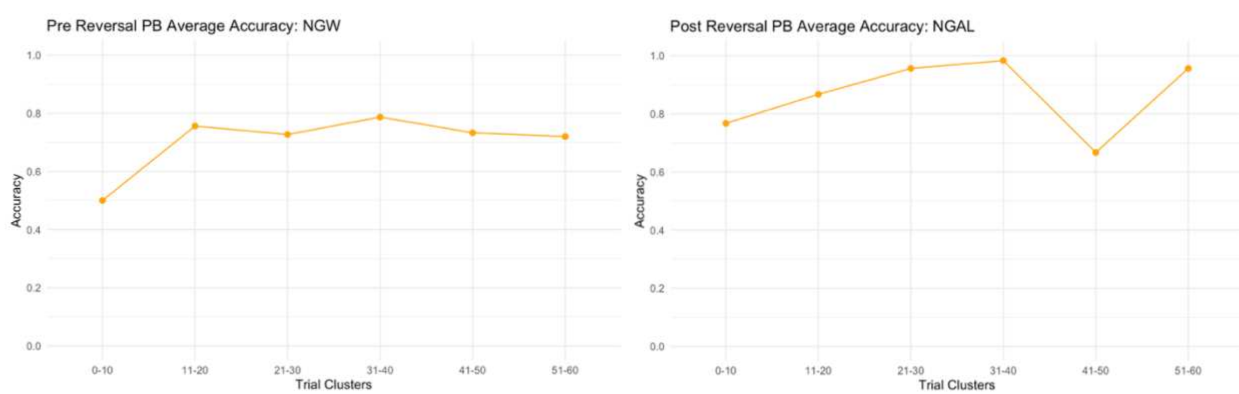


Figure 26

Pre-reversal: Pre-reversal accuracy in NGW started off around chance but quickly gained a higher average accuracy. Then throughout block 3 we see a steady accuracy that lingered around 80% which is remarkable as in Experiment 1, NGW pre reversal was always below chance level. Due to an overpowering Pavlovian effect, we do see the accuracy not reaching 100% as a subgroup of participants potentially did not learn the association of the valence and action due to the incongruency, refer to figure 22 to note the improved learning for NGW over time.

Post-reversal: Post reversal ‘No Go to Win’ changed to ‘No Go Avoid Losing’, a compatible response in which the instrumental learning ‘no go’ and Pavlovian bias ‘avoid losing’ are aligned. Participants' accuracy significantly increased post reversal which highlights their ability to learn instrumental response of suppressing action but can also be ascribed to modified valence.

No Go to Avoid Loss to No Go to Win

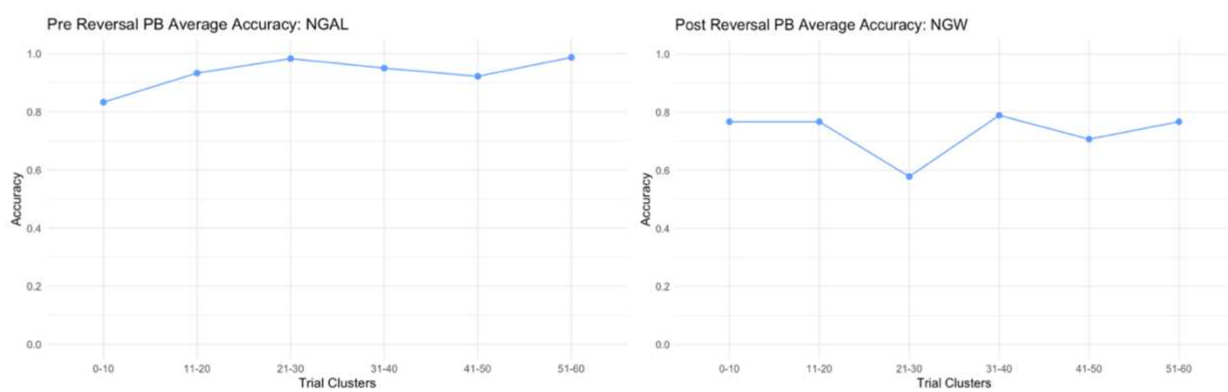


Figure 27

Pre-reversal: The ‘No-Go to Avoid Losing’ trial type average accuracy remained constantly high and near ideal performance levels throughout the pre-reversal phase. NGAL is a congruent trial type which led to high accuracy where the valence of avoiding loss assisted in suppressing action. The inhibition of instrumental response leads to avoiding loss for this condition.

Post-reversal: Post reversal, 'No Go to Avoid Loss' shifted to 'No Go to Win'. Instead of avoiding a loss, participants were now receiving a reward, which did not align with their instrumental response of inaction. Average accuracy decreased when participants were presented with the reversed contingencies to around 70%, fluctuating throughout block 4. Participants clearly found it difficult to refrain from acting in order to receive a reward, which goes against participants' instinctive desires to act in order to receive rewards, is likely the cause of this decline. Regardless, we still see that participants were able to switch the responses rapidly in the face of reversal which we did not predict would occur due to the rigid nature of Pavlovian biases. Additionally, we note that the change from NGAL to NGW was salient from the first trial cluster, indicative of participants ability to learn the new Pavlovian relationship rapidly.

Discussion on Reversal of Pavlovian Biases

With the introduction of reversal and modified contingencies we saw a strong Pavlovian influence within participants' performance. Their prominence was especially salient when the trial type was incongruent, meaning that the valence and action were misaligned, e.g. NGW. Participants showed a strong tendency to 'go' with a 'win' bias and withhold their response during a 'avoid loss' bias. Moreover, the biases were more pronounced in the 'go' circumstances (GW & GAL) than in the 'no-go' conditions (NGW & NGAL), indicating that the Pavlovian extensions might be more salient and overpowered when it comes to action over inaction. In most cases in the reversal of Pavlovian bias, we noticed that in the midst of post reversal, participants realized that the valence had changed and reacted with either suppression or invigoration of action, leading to fluctuations within the trials. For example, accuracy decreased significantly after reversal in the GW to GAL condition, indicating a discrepancy between the initial instrumental response and the updated Pavlovian connection however was rapidly increased once participants realized that a change in outcome had occurred. This was in line with our hypothesis as reversing the valence was hard for participants, resulting in decreased accuracy however we did see some flexibility and the employment of declarative knowledge which helped participants regain their lost accuracy. For example, NGW reversing to NGAL (Figure 25) saw a significant increase in the average accuracy, indicative of participants cognitive flexibility. These biases did show differences during reversal when compared to goal-directed instrumental learning however, they also showed ability to rapidly reverse contingencies by employing explicit knowledge of the changes made within outcomes post reversal.

Evaluation of Post-Task Surveys

In Experiment 2, one of the core modifications was the addition of a post-task survey. This survey's overarching purpose was to analyze if participants were gaining explicit knowledge of the associations presented to them in the new deterministic and an extra block paradigm and additionally, able to verbally describe the strategy they utilized to win rewards as well as avoid losses. The surveys assisted us in capturing idiosyncratic patterns for individual variability that participants demonstrated on the reversals of the two conditions: Pavlovian bias and goal-directed instrumental learning. The survey asked questions that assessed participants ability to state the valence and action for the four robots (refer to Appendix). For the two reversal conditions of goal-directed instrumental learning and Pavlovian biases, we had twice the number of questions for our participants – the first set asking about the pre reversal associations and the second inquiring about the correct associations post reversal. This assisted us in parsing out how participants were performing in each of the three conditions and compare them amongst each other (refer to figure 27).

Post-Task Survey Results

After successfully collecting all of our 90 participants' data for the post-task survey, we analyzed it utilizing R Studio by creating visualizations. Due to concerns about the quality of answers written by certain participants, we had to discard eight post-task surveys. This led us to evaluate a total of 82 surveys for this section of the assessment.

Baseline (2a: No Reversal Condition)

To begin, the accuracy of participants in baseline, no reversal condition must be assessed. Table 1 shows the average accuracy of performance on the four trial types (GW, GAL, NGW, NGAL) in different contexts of valence and action. The values represented in Table 1 represent the percentage of participants that correctly answered the valence or action for each trial type in the post-task survey. ‘Valence’ refers to the outcome associated with one of the four trial types (gain or loss). Conversely, ‘Action’ refers to the action associated with the particular trial type (go or no-go). Table 1 also notes a value for ‘valence_perfect’ where we calculated the average number of participants with perfect scores (all 4 robots had valence correctly identified) across all participants. This stays true for the calculation of ‘action_perfect’ as well, where we calculated the percentage of participants who were perfect on the action subsection of the post-task survey.

Table 1.

Metric	Average
Valence_NGW	0.806
Action_NGW	0.774
Valence_GW	1.000
Action_GW	0.968
Valence_GAL	0.710
Action_GAL	0.967
Valence_NGAL	0.774
Action_NGAL	0.645
Valence_Perfect	0.483
Action_Perfect	0.645

**Note. Accuracy values are calculated as the proportion of correct responses.

Table 1. Mean Accuracy for Valence and Action Across Conditions. The table represents the average performance in each trial type. Note that 'perfect' was not a trial type and was calculated separately after the post-task survey to assess participants who were able to answer all the questions correctly.

Table 1 demonstrates what conditions led to the greatest amount of explicit, verbalizable knowledge. By assessing the average accuracy of performance, we can see that participants scored the highest on the 'Go to Win' condition, as shown by the very high scores (Valence = 1.00; Action = 0.97). This is congruent with performance on the go/no-go task (compare with Figure 15). Additionally, another salient result we found was the decreased accuracy of trial types of GAL and NGW in both action as well as valence. This suggests that even with four blocks of training these trial types posed a challenge to participants potentially due to the misalignment in Pavlovian tendencies that also resulted in impaired performance in these conditions in the go/no-go task. We calculated the average valence score across all four trial types (Table 1) which resulted in the value, 0.483 vs. the average action scores across the trial types which came out to be 0.645. The trends in the average accuracy in post-task survey remain consistent with what we observed in our data for average accuracy of each trial type in Experiment 2a (Figure 15), indicating that overall, the patterns we see in the go/no-go task are consistent with participants' explicit knowledge of the task. For example, performance on NGW was low in the go/no-go task, and participants showed lower accuracy in reporting the valence and action for NGW in the questionnaire.

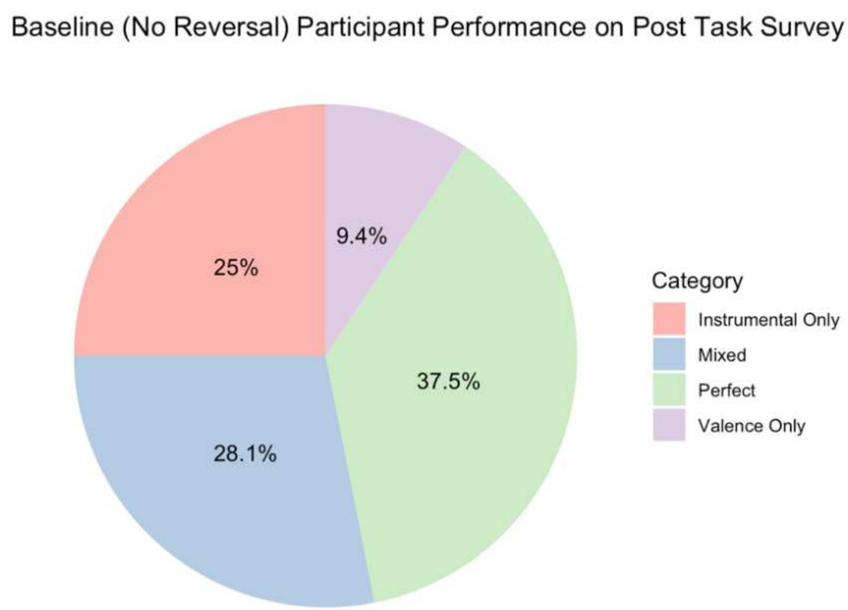


Figure 28: Pie chart represents participant categorization in the baseline condition, no reversal condition.

We noted a lot of heterogeneity in performance in the go/no-go task, with some participants performing at high levels across all conditions without showing a clear Pavlovian bias, and others showing performance at chance or even reversed (which can happen if the participant responds only on the dimension of valence in the incongruent conditions NGW and GAL). In order to see if there were similar heterogeneity in the post-task questionnaire, and to relate participants' explicit knowledge to their task performance, we classified participants into four categories based on the post-task questionnaire, as illustrated in Figure 28. In Figure 28, instrumental only refers to participants who performed with a perfect accuracy (100%) on the post-task survey questions evaluating the instrumental response associated with each robot, but who did not answer all the questions about valence correctly. Similarly, valence only refers to the percentage of participants that were able to answer questions about the outcome of the association correctly – with a 100% accuracy for all four trial types, but did not answer all the questions about the instrumental response correctly. Participants who answered all eight questions correctly including both instrumental response and valence were categorized as 'perfect', contrarily participants who were not able to perform with a perfect accuracy for either valence or action, were binned into the 'mixed' category. With the results shown in Figure 26, we see a large proportion of participants, ~38% who were able to accurately answer all eight questions regarding the valence and action of each of the four trial types, this suggests that a large subgroup of the population was able to learn all the associations well and performed perfectly on the task. Next, we see a substantial percentage in the mixed category (28%) which tells us that more a quarter of participants did not have a clear mastery of either the valence or instrumental response throughout the task. Then we see 25% participants answer accurately only the questions regarding the instrumental response and 9% answered accurately only the questions about the valence or outcome of each trial type. This leads us to infer that instrumental conditions might be easier for the participants to learn explicitly as these were controllable actions. We could argue that participants who ignored the 'valence' were being rational as the actions they performed should have taken precedence on their learning strategy. Additionally, the high proportion of 'mixed' individuals could be explored further to understand individual differences, i.e. cognitive flexibility or attention spans.

Instrumental Learning Reversal (2b)

Pre reversal post-task survey: When comparing the pre reversal average accuracy in the instrumental reversal condition, we see parallels with the baseline condition (Table 1). There is a similarity of pattern of averages where, in the valence, GW (.929) was performed the best and NGW (.500) the worst. Meanwhile, in the action subgroup, we saw GW perform the best (0.893) while NGW (0.571) the worst. There are a few discrepancies but those could be attributed to two things. Firstly, participants did not have as much training as they did in the baseline condition (240 trials > 180 trials). Recall that in block 4 in the instrumental condition, we reverse contingencies therefore participants did not have an additional 60 trials that they had in the baseline condition. Additionally, we can expect some interference in memory due to intrusions from the post-reversal task.

Post reversal post-task survey: For the mean accuracy for post-task, after the reversal of action contingencies (Go became No Go and vice versa) we saw a collective decrease in the average accuracy. GW for the valence subgroup, where participants had to answer questions based on the outcome of the stimuli, performed the best (.821) whereas GAL was the worst (0.607). When considering the performance in action group, where participants had to correctly identify the actions associated with a trial type, NGAL performed the best (0.964) and NGW & GAL the worst (0.750). We can hypothesize that GAL performed the worst due to the shift from NGAL to GAL in the fourth block which could have interfered with the memory participants had for what they had initially learned in the task.

Table 2. Instrumental learning (Pre reversal): Mean Accuracy for Valence and Action Across Conditions

Metric	Average
Pre_Valence_NGW	0.500
Pre_Valence_GW	0.929
Pre_Valence_GAL	0.679
Pre_Valence_NGAL	0.679
Pre_Valence_Perfect	0.321
Pre_Action_GAL	0.642
Pre_Action_GW	0.893
Pre_Action_NGAL	0.857
Pre_Action_NGW	0.571
Pre_Action_Perfect	0.535

Note. Accuracy values represent the proportion of correct responses for each condition.

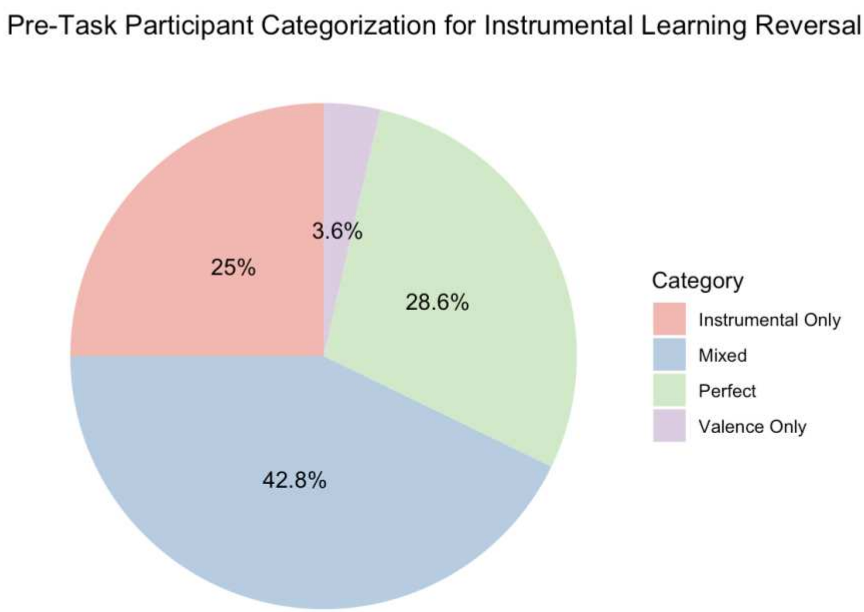
Table 3 Instrumental learning (Post reversal): Mean Accuracy for Valence and Action Across Conditions

Metric	Average
Post_Valence_NGW	0.714
Post_Valence_GW	0.821
Post_Valence_GAL	0.607
Post_Valence_NGAL	0.786
Post_Valence_Perfect	0.464
Post_Action_GAL	0.750
Post_Action_GW	0.893
Post_Action_NGAL	0.964

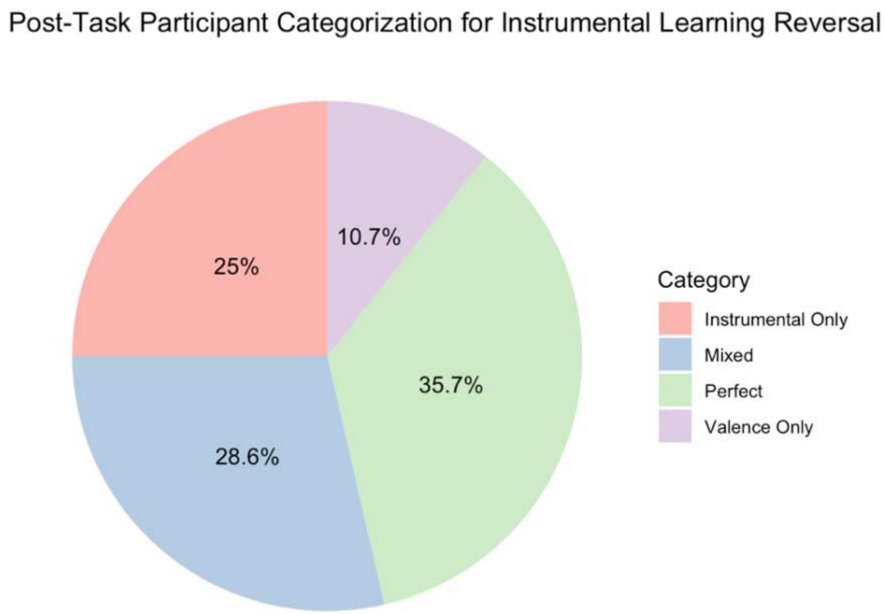
Metric	Average
Post_Action_NGW	0.750
Post_Action_Perfect	0.607

Note. Accuracy values represent the proportion of correct responses for each condition.

Below, figure 29 represents the categories participants were placed in depending on how they performed in the post-task survey (refer to criteria described in baseline performance).



A.



B.

Figure 29: Pie chart represents participant categorization in the pre and post instrumental learning condition.

In the pre reversal condition (Figure 29 A) we see the majority of participants in the mixed category (~43%), recall that in baseline the mixed category was approx. 28%, indicative of their inability to learn the associations completely for the trial types prior to the reversal. This could be due to the fact that participants did not have an extra block (similar to baseline) to train more and understand the associations. However, we still see a great proportion of participants who were in the perfect category, ~29%. A noteworthy observation was the percentage of participants that learned valence only (~4%), which could have been counterproductive when the contingencies changed. Instrumental response only was at 25%, which could be due to the ease at which actions are learned and employed with practice.

Pavlovian Bias Reversal (2c)

Pre reversal post-task survey: When comparing the pre reversal average accuracy for Pavlovian condition, we observe similarities with the baseline condition (Table 1). As was the case for instrumental reversal, there were some differences which can be attributed to the decreased number of training blocks (240 trials compared to 180 trials). Regardless, we do see a similar pattern of averages where for valence, GW (.828) performed the best and GAL & NGAL (.690) the worst. Meanwhile, for action, we saw NGAL perform the best (0.862) while NGW (.655) was the worst.

Table 4. Pavlovian learning (Pre reversal): Mean Accuracy for Valence and Action Across Conditions

Metric	Average
Pre_Valence_NGW	0.724
Pre_Valence_GW	0.828
Pre_Valence_GAL	0.690
Pre_Valence_NGAL	0.690
Pre_Valence_Perfect	0.414
Pre_Action_GW	0.793
Pre_Action_GAL	0.724
Pre_Action_NGW	0.655
Pre_Action_NGAL	0.862
Pre_Action_Perfect	0.551

Note. Accuracy values are calculated as the proportion of correct responses.

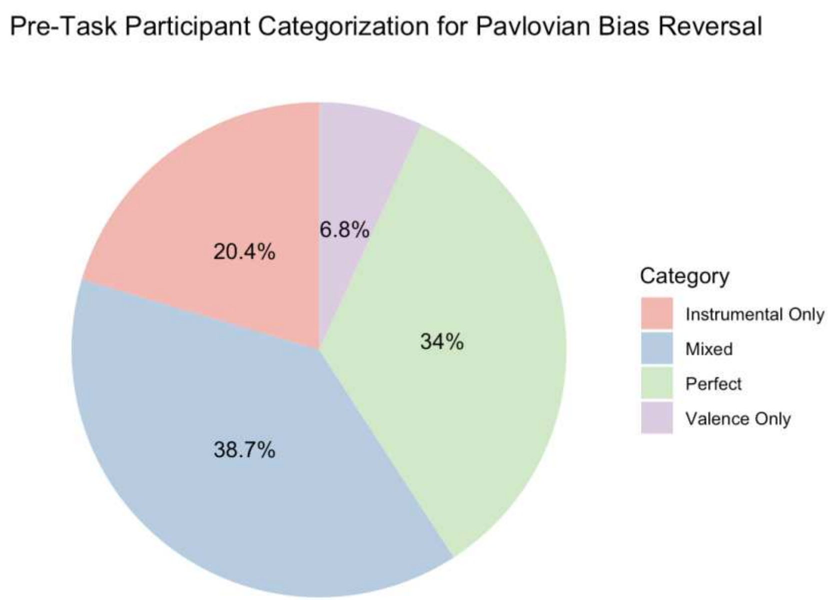
Table 5. Pavlovian learning (Post reversal): Mean Accuracy for Valence and Action Across Conditions

Metric	Average
Post_Valence_NGAL	0.655
Post_Valence_GAL	0.482
Post_Valence_GW	0.621
Post_Valence_NGW	0.552
Post_Action_GW	0.965
Post_Action_GAL	0.724
Post_Action_NGW	0.758
Post_Action_NGAL	0.931
Post_Valence_Perfect	0.414
Post_Action_Perfect	0.655

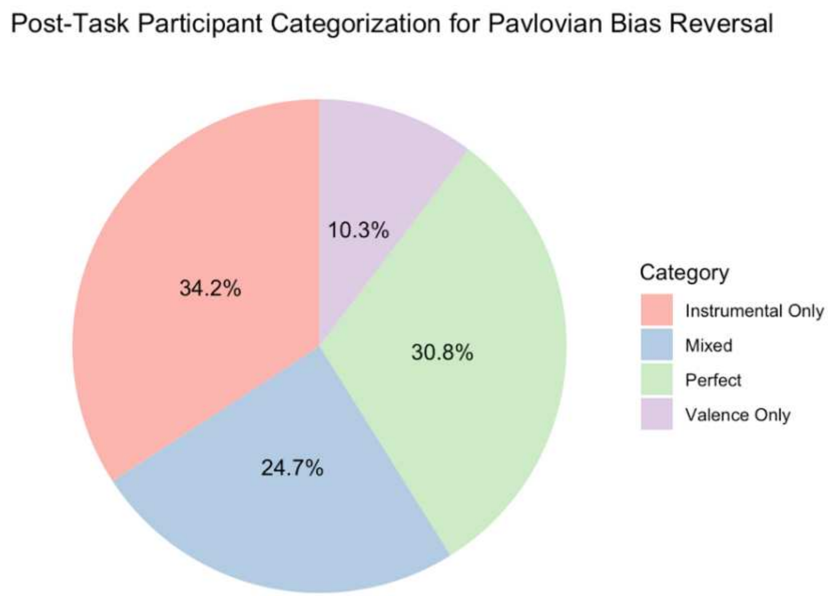
Note. Accuracy values are calculated as the proportion of correct responses.

Post reversal post-task survey: For the mean accuracy for post-task survey, after the reversal of valence contingencies ('win' became 'avoid loss' and vice versa) we saw a collective decrease in the average accuracy (Table 5). NGAL in valence subgroup, where participants had to answer questions based on the outcome of the stimuli, performed the best (.655) whereas GAL the worst (0.482). When considering the performance in action group, where participants had to correctly identify the actions associated with a trial type which remained consistent pre and post reversal in the Pavlovian condition, GW performed the best (0.965) and GAL the worst (0.724). In this instance, we can hypothesize that GAL performed the worst due to the shift from GW to GAL in the fourth block which could have created a challenge for participants from what they had initially learned in the task. Note that post reversal in instrumental condition (Table 3) has higher average accuracy in all accounts than post reversal Pavlovian condition (Table 5). Additionally, we also see that participants were unable to explicitly learn the associations in the Pavlovian condition to perform well post reversal leading to a significant decrease after the reversal was implemented.

Below, Figure 30 shows the four aforementioned categories of instrumental only (refers to participants who performed with a perfect accuracy), valence only (refers to the percentage of participants that were able to answer questions out the outcome of the association correctly), 'perfect' criteria (participants who answered all eight questions correctly) and lastly, 'mixed' category where participants who were not able to perform with a perfect accuracy for either valence or action.



A.



B.

Figure 30: Pie chart represents participant categorization in the pre and post instrumental learning condition.

With the results shown in Figure 30A, we see a large proportion of participants, ~39% who were not able to accurately answer all eight questions regarding the valence and action of each of the four trial types which suggests overreliance on either valance or action. Perfect category (34%) demonstrates that a large subgroup of the population was able to learn the associations well and performed perfectly on the task. In participants who were not perfect overall but performed with 100% accuracy in one dimension, we see 20% participants answer accurately the questions regarding the action of each trial type or the instrumental response (going or suppressing action) and lastly only 7% were able to correctly answer the valence or outcome of each trial type.

After implementing reversal (Figure 30B) we observe an increase in participants' accuracy on instrumental and valence conditions. Mixed category (~24%) decreased post reversal and a larger proportion of participants were now either completely accurate on action questions or valence questions. We could infer that the increase in valence is due to the fact that post reversal valence was switched which led to participants focusing more on how the previous trial type was now associated with what outcome. Similarly, instrumental response could have increased due to increased amount of practice as the instrumental response for each of the four trial types remained the same, participants did not have to adjust their actions which led them to solidify their explicit information on actions. Lastly, we could also attribute the performance to participants showing a recency bias and interference. Overall, we see the rapid adjustment that participants made after the reversal occurred, leading to a higher proportion of participants in the valence, instrumental and perfect categories. To support what we found in our analysis, we visually depicted how participants grouped in certain categories performed in the actual task (Figure 30).

Explicit vs Implicit Information Transfer

In the following section, we plotted participants scores from Experiment 2 go/no-go task to compare their performance on post-task survey. We then divided their post-task survey performance into 4 graphs based on their survey performance namely: perfect, valence perfect only, action perfect only and lastly, mixed performance where the participants neither answered the questions perfectly about action or valence. In baseline condition (Figure 31 & Figure 32), we used block 4 from the task performance as there were no reversals involved and created 4 graphs from the categories mentioned above based on their survey performance. However, for our reversal conditions (instrumental learning and Pavlovian biases) we used pre reversal block 3 and post reversal block 4 from the go/no-go task performance and created eight separate graphs for post-task performance, four for the pre reversal post-task accuracy and four for the post reversal post-task accuracy (refer to Appendix B). The aim of this section was to evaluate each participant's ability to explicitly state the contingencies (especially for different types of explicit knowledge regarding action, valence or both) and compare their explicit learning to their accuracy levels in the actual task.

Baseline

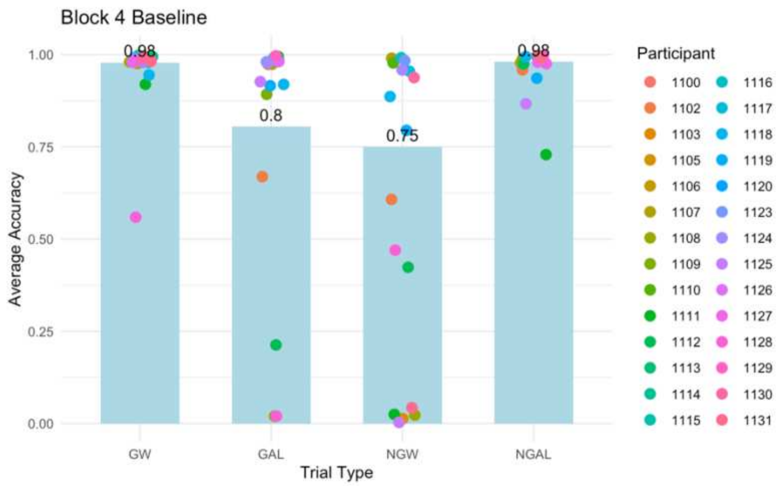
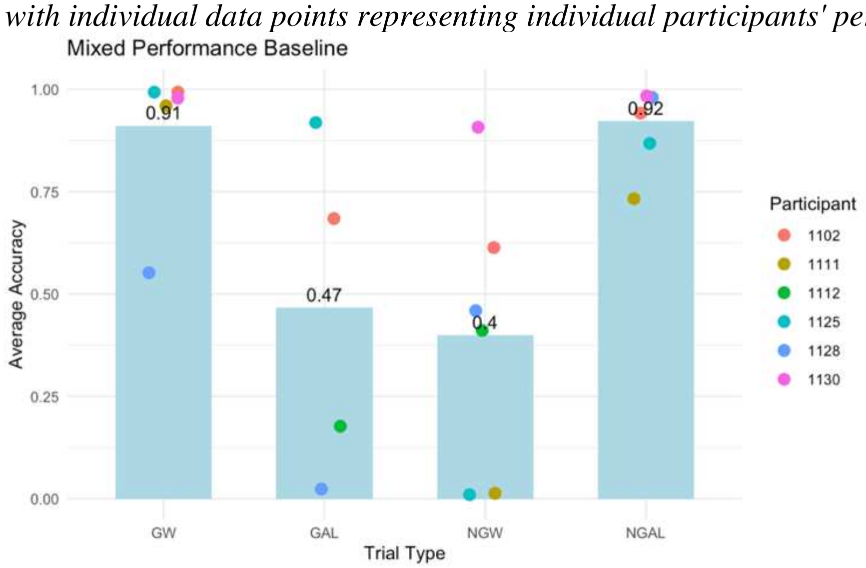
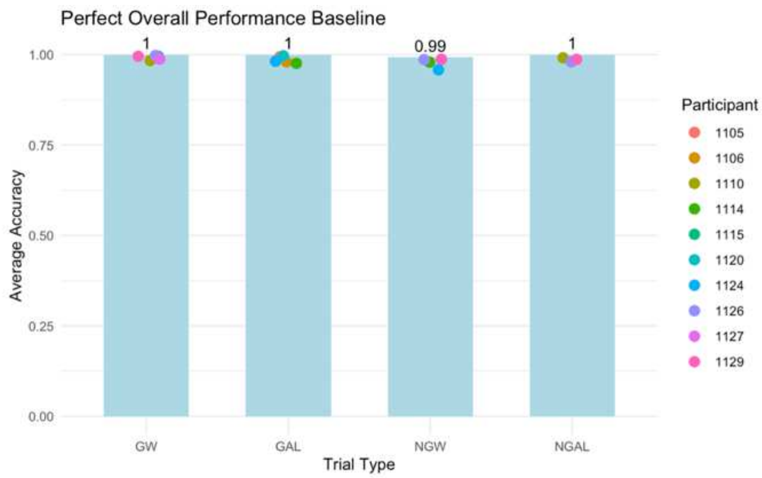


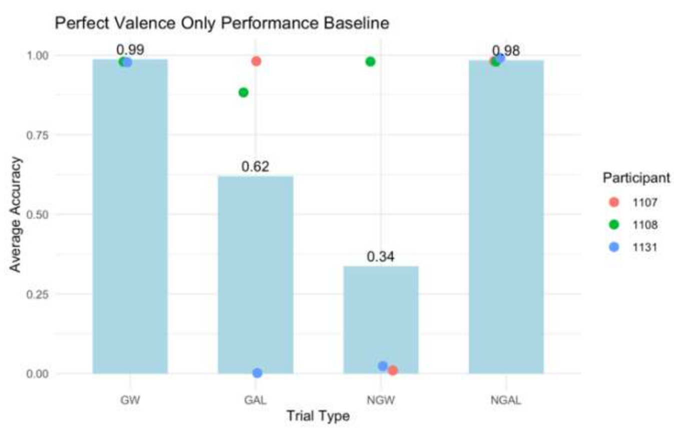
Figure 31: Average accuracy performed on each trial type across all blocks in Experiment 2a (baseline) with individual data points representing individual participants' performance.



A


Perfect Overall Performance Baseline
1.00
0.75
0.50
0.25
0.00
Average Accuracy
1
1
0.99
1
GW
GAL
NGW
NGAL
Trial Type
Participant
1105
1106
1110
1114
1115
1120
1124
1126
1127
1129

B.


Perfect Valence Only Performance Baseline
1.00
0.75
0.50
0.25
0.00
Average Accuracy
0.99
0.62
0.34
0.98
GW
GAL
NGW
NGAL
Trial Type
Participant
1107
1108
1131

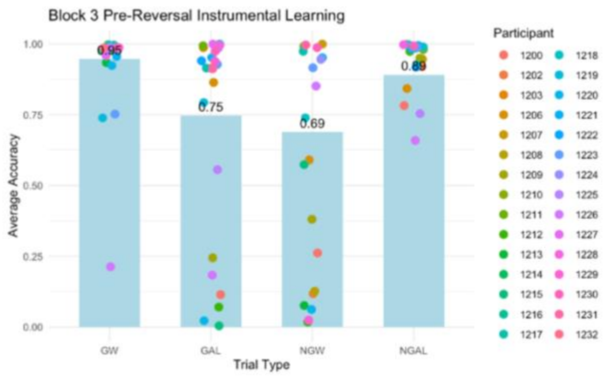
C.


Perfect Action Only Performance Baseline
1.00
0.75
0.50
0.25
0.00
Average Accuracy
0.99
0.98
0.82
0.99
GW
GAL
NGW
NGAL
Trial Type
Participant
1100
1103
1113
1116
1117
1118
1119
1123

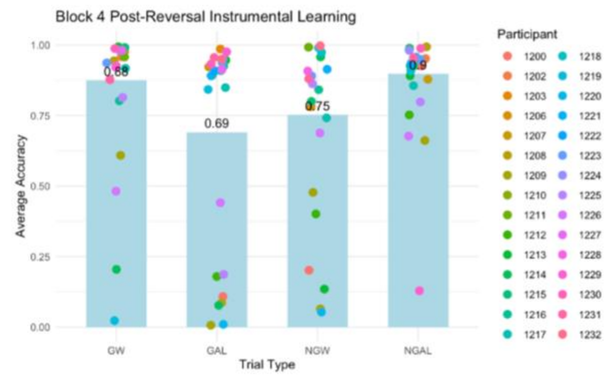
D.

Figure 32: Specific participants categorized in groups of their optimal performances in post-task survey. A. represents participants who were unable to score perfectly on both valence or action; B. represents participants who were able to score perfectly on both valence and action in the post-task survey; C. represents individuals who were only able to score perfectly in valence condition while D. represents individuals who were only able to obtain a perfect score in action section of the post-task survey.

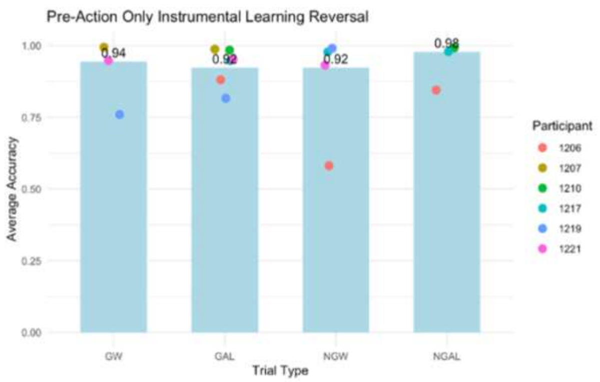
After evaluating the mixed condition (Figure 32A) we notice an overlap between participants who were categorized under 'mixed' due to their inability to capture either action or valence and their lower average accuracy in block 4 of the task (Figure 31). For example, participant 1102 – among others – was unable to score in the NGW condition even after extensive training, which was evident in their performance on the post-task survey as well. Participants in 'Perfect' overall condition (Figure 32B) were able to perform with high average accuracy for all four of the trial types – as expected. Additionally, participants in the 'perfect action' category showed high average accuracy for all four trial types. However, those in the 'valence' category showed some struggle performing on incongruent trial types such as NGW and GAL. For example, 1131 (perfect in valence condition) struggled to maintain a high accuracy in NGW condition potentially due to overreliance on reward as an outcome, that could have invigorated the participant.


Block 3 Pre-Reversal Instrumental Learning
Average Accuracy
Trial Type
GW
GAL
NGW
NGAL
0.95
0.75
0.69
0.89
Participant
1200 1202 1203 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217
1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232

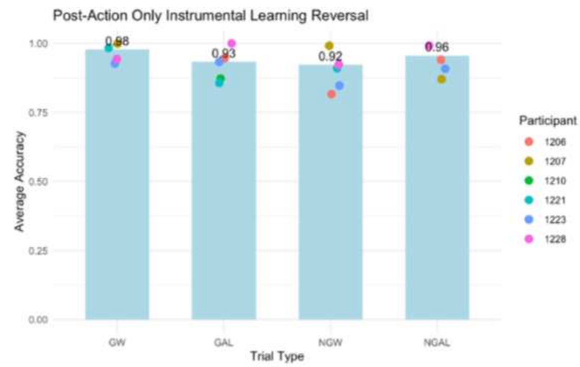
A.


Block 4 Post-Reversal Instrumental Learning
Average Accuracy
Trial Type
GW
GAL
NGW
NGAL
0.88
0.69
0.75
0.9
Participant
1200 1202 1203 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217
1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232

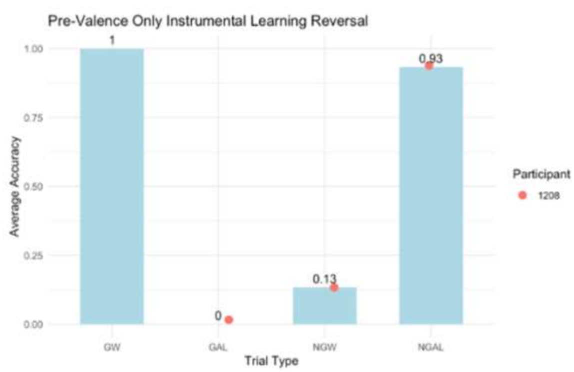
B.


Pre-Action Only Instrumental Learning Reversal
Average Accuracy
Trial Type
GW
GAL
NGW
NGAL
0.94
0.92
0.92
0.98
Participant
1206
1207
1210
1217
1219
1221

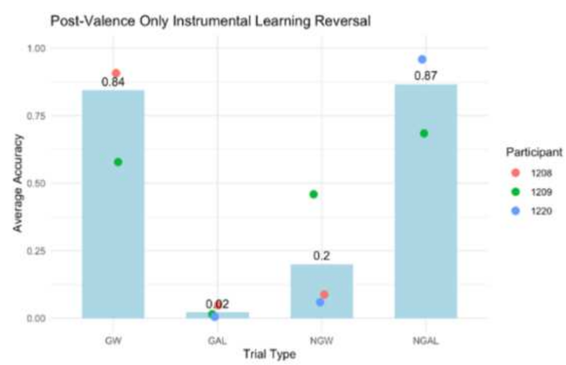
C.


Post-Action Only Instrumental Learning Reversal
Average Accuracy
Trial Type
GW
GAL
NGW
NGAL
0.98
0.93
0.92
0.96
Participant
1206
1207
1210
1221
1223
1228

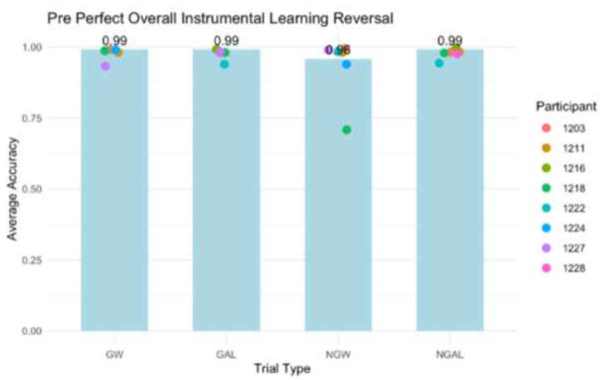
D.


Pre-Valence Only Instrumental Learning Reversal
Average Accuracy
Trial Type
GW
GAL
NGW
NGAL
1
0
0.13
0.93
Participant
1208

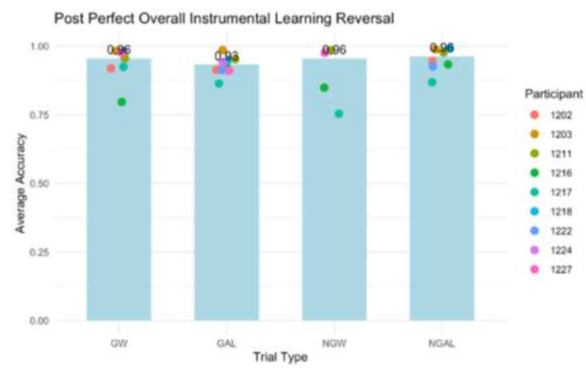
E.


Post-Valence Only Instrumental Learning Reversal
Average Accuracy
Trial Type
GW
GAL
NGW
NGAL
0.84
0.02
0.2
0.87
Participant
1208
1209
1220

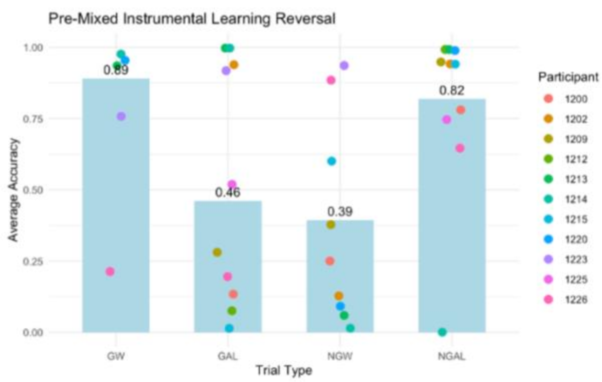
F.



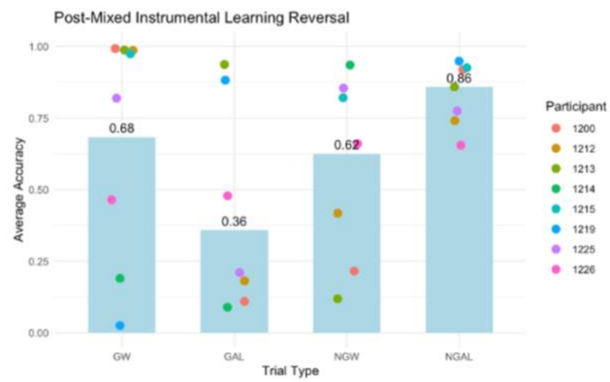
G.



H.



I.

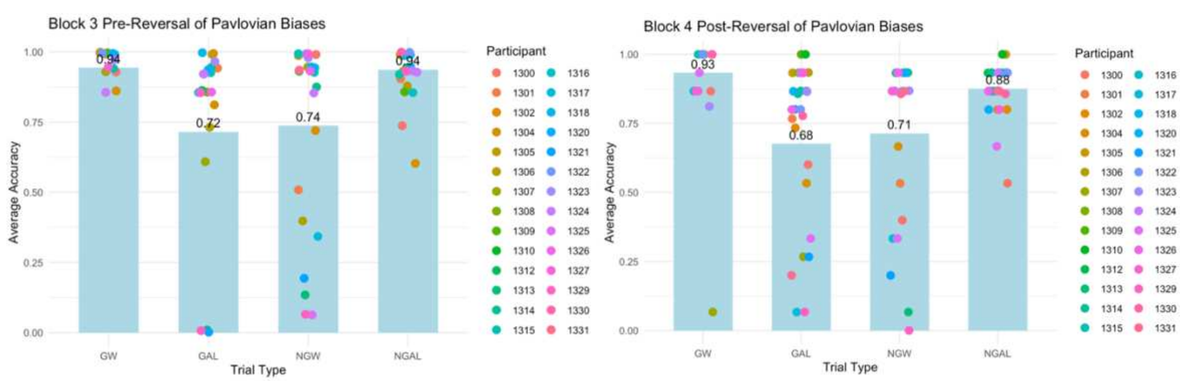


J.

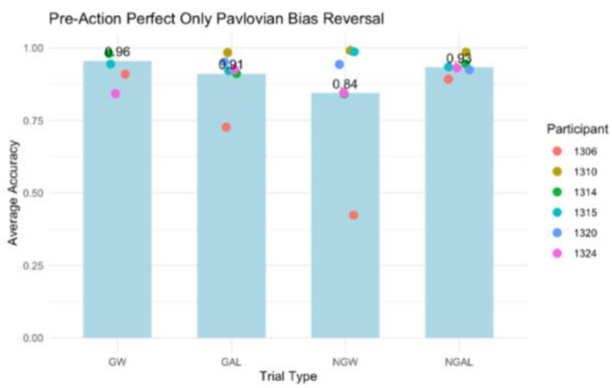
Figure 33: Specific participants categorized in groups of their optimal performances. A & B represents pre and post reversal task performance; C & D represent participants accuracy on the post-task survey regarding the action related to the stimuli, specifically, C was for pre reversal and D for post reversal; E & F represent pre and post valence, respectively, for participants who were able to score perfectly only in the valence condition; G & H represent pre and post perfect overall depiction of participants who were able to not only attain a perfect score for valence but also action. Lastly, I & J represents the 'mixed' categories where participants listed did not score perfectly on either valence or action, both pre and post reversal questions on the post-task survey.

Figure 33 provides an overview of the relationship between participants ability to verbalize the associations of specific stimuli, both pre and post reversal, and their performance on the go/no-go task conditions. These graphs demonstrate participants ability to use explicit knowledge to achieve optimal learning as seen by their high average accuracy on the task for perfect overall, perfect action and mixed figures. For example, participant 1210 was able to score perfectly on the task post reversal and also was able to transfer that knowledge on post reversal action association in the post-task survey for the robots shown. Similarly, in pre and post perfect category in the survey, participant 1224 was able to score perfectly, which was indicative of their performance which was above average in the actual task, before and after reversal, as expected. However, in alignment with what we observed in baseline condition, participants who scored correctly on valence only in the post-task survey, struggled to demonstrate high average accuracy in the task, especially on trial types that were incongruent (NGW, GAL). Lastly, participants in the mixed condition were not able to optimize their performance either with a subset of population within this category struggling to maintain a high average accuracy. Overall, we see a trend of valence and mixed category participants who were unable to perform optimally on the task while participants in perfect action and perfect overall categories performed superiorly.

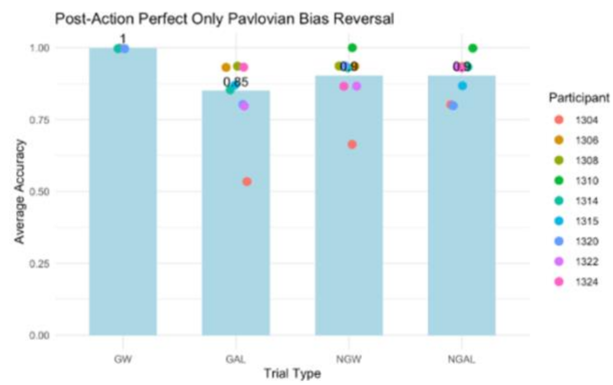
Reversal of Pavlovian Bias



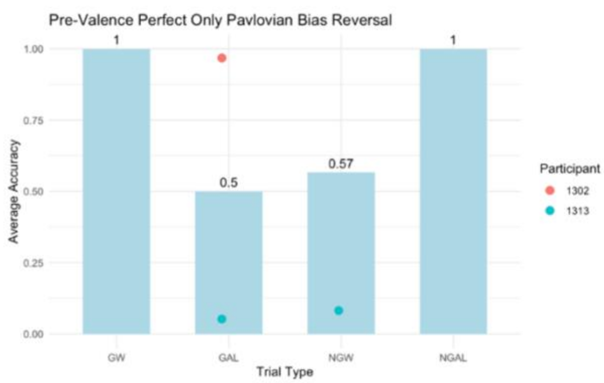
A B


Pre-Action Perfect Only Pavlovian Bias Reversal
Average Accuracy
1.00
0.75
0.50
0.25
0.00
0.96
0.91
0.84
0.93
GW
GAL
NGW
NGAL
Trial Type
Participant
1306
1310
1314
1315
1320
1324

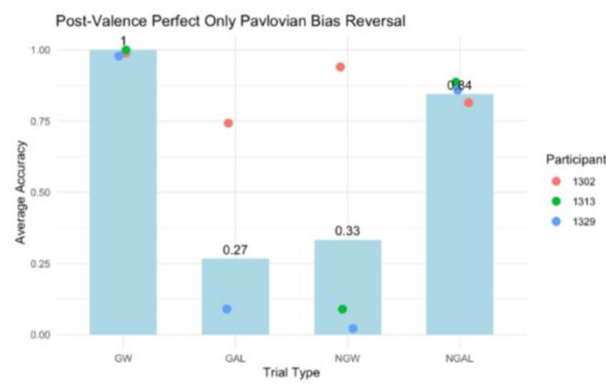
C.


Post-Action Perfect Only Pavlovian Bias Reversal
Average Accuracy
1.00
0.75
0.50
0.25
0.00
1
0.85
0.9
0.9
GW
GAL
NGW
NGAL
Trial Type
Participant
1304
1306
1308
1310
1314
1315
1320
1322
1324

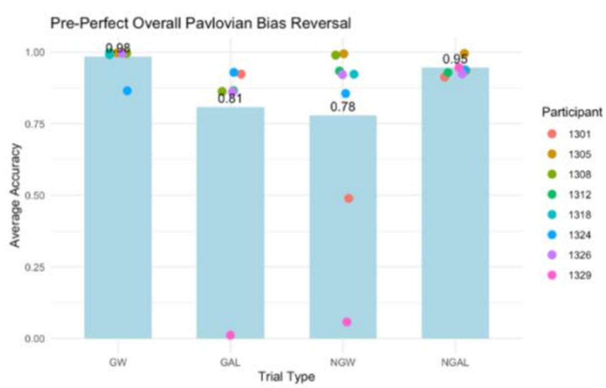
D.


Pre-Valence Perfect Only Pavlovian Bias Reversal
Average Accuracy
1.00
0.75
0.50
0.25
0.00
1
0.5
0.57
1
GW
GAL
NGW
NGAL
Trial Type
Participant
1302
1313

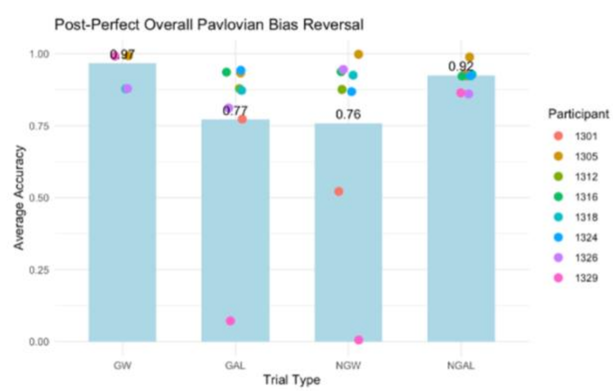
E.


Post-Valence Perfect Only Pavlovian Bias Reversal
Average Accuracy
1.00
0.75
0.50
0.25
0.00
1
0.27
0.33
0.84
GW
GAL
NGW
NGAL
Trial Type
Participant
1302
1313
1329

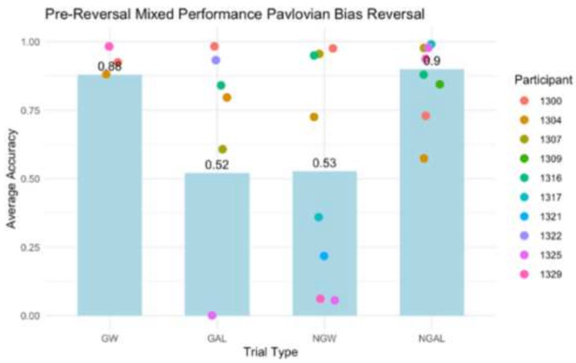
F.


Pre-Perfect Overall Pavlovian Bias Reversal
Average Accuracy
1.00
0.75
0.50
0.25
0.00
0.98
0.81
0.78
0.95
GW
GAL
NGW
NGAL
Trial Type
Participant
1301
1305
1308
1312
1318
1324
1326
1329

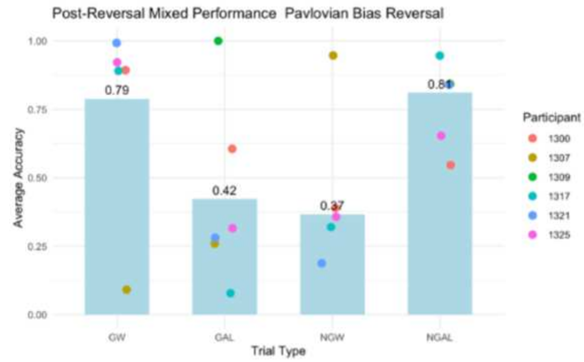
G.


Post-Perfect Overall Pavlovian Bias Reversal
Average Accuracy
1.00
0.75
0.50
0.25
0.00
0.97
0.77
0.76
0.92
GW
GAL
NGW
NGAL
Trial Type
Participant
1301
1305
1312
1316
1318
1324
1326
1329

H.



I.



J.

Figure 34: Specific participants categorized in groups of their optimal performances. A& B represents pre and post reversal task performance; C & D represent participants accuracy on the post-task survey regarding the action related to the stimuli, specifically, C was for pre reversal and D for post reversal; E & F represent pre and post valence, respectively, for participants who were able to score perfectly only in the valence condition; G & H represent pre and post perfect overall depiction of participants who were able to not only attain a perfect score for valence but also action. Lastly, I & J represents the 'mixed' categories where participants listed did not score perfectly on either valence or action, both pre and post reversal questions on the post-task survey.

Similar to previous findings, we encounter participants ability to explicitly verbalize the stimuli association presented in the task. Participant categorized as 'perfect action only' performed accurately for all four trial types, relying on their ability to consciously recall action associated with the presented stimuli. For example. participant 1320 – amongst others – was able to correctly identify the robots' actions both pre and post reversal. Participants who were able to correctly score on both the valence and action for the post-task survey, under 'perfect overall' category such as 1329, performed correctly and above average on the task itself. However, individuals who performed perfectly on valence in post-task survey found it challenging to retain their performance above chance, specifically for trial types where actions did not align with valence. For example, participant 1302 who was able to correctly identify the valence of each trial type before and after reversal seemed to find it challenging to maintain their accuracy high post reversal in the task itself – specifically in trial type NGW and GAL. Similarly, for mixed category, participants mostly struggled with the two incongruent trial types – GAL and NGW. Participant 1317 who was in the mixed category performed around chance for both pre and post reversal, especially on GAL and NGW.

Discussion Experiment 2

The modified paradigm we created for Experiment 2 served as a better indicator of how Pavlovian biases and goal-directed instrumental learning interacted across the three conditions, namely: baseline (2a), reversal of goal-directed instrumental learning (2b), and reversal of Pavlovian bias (2c). We identified consistent patterns of Pavlovian bias and how they facilitate and hinder goal-oriented actions. There were four key findings that we found with Experiment 2. Firstly, we observed the superior performance of trial types of GW and NGAL in all three experiments, regardless of the reversal, both of these trial types persevered with the highest proportion of average accuracy. This means that when congruent with goal-directed behaviors, Pavlovian biases act as catalysts to the required environment, e.g. these biases will invigorate or suppress a response in alignment to our goal-directed or conscious actions when faced with a familiar stimulus. Conversely, when faced with situations where our actions are not congruent with the valence present, these biases often interfere with our actions, meaning that when participants were faced with stimuli where you had to take action 'Go' to 'avoid a loss' they had trouble aligning these two contingencies together, leading to decreased accuracy overall. Secondly, within the reversal of Pavlovian biases (Experiment 2c), we did see rigidity however it was not as impairing as we had initially posited. Performance of participants post reversal of Pavlovian biases (Figure 23) did not go below chance (50%), meaning that participants had learned the switch of valence rapidly and were able to apply their goal-directed behaviors to score a higher average accuracy than chance. This may be due to the deterministic nature of the modified paradigm which could be one of the limitations of the general go/no-go task that could be addressed. Therefore, we found that Pavlovian biases do indeed exhibit flexibility – specifically in the present study. Thirdly, we noticed that some participants focused solely on the valence, especially in incongruent trial types NGW and GAL (Figure 16 & Figure 22). These individual differences require a separate stream of conversation to discuss how salient valence values like going or suppressing action could play a crucial role in dictating conscious action. Lastly, there was a persistent general trend of ‘go’ biases throughout Experiment 2, as evidenced by stronger learning in the ‘go to win’ condition compared to the ‘no-go to avoid losing’ condition. Research suggests that when one engages in action for decision making individuals feel more ‘active’ and in control than when they are required to suppress action for an optimal response (Rosenblatt, 2003). Such patterns can also be seen within our task where participants were relying on action more than inaction, even if suppressing action led to a better outcome. This bias highlights the tendency for go actions to be more readily learned and reinforced than no-go actions.

General Discussion

In the current study, we examined Pavlovian biases and their interaction with goal-directed instrumental learning. The overarching aim of this study was to evaluate the rigidity Pavlovian biases exhibit, especially in the face of change. Previous research has argued that Pavlovian biases are a decision-making phenomenon where a neutral stimulus – after being continuously paired with a certain positive or negative valence – gains a predictive and motivational importance (Berridge et al., 2000; Robinson & Flagel, 2009). As a result of the association, the environmental cue attains the ability to influence choice due to its predictive power. These biases can have powerful influence over goal-directed actions (Clark et al., 2012).

In this study, we explored the interaction of Pavlovian biases with goal-directed instrumental learning, which is a type of learning where actions are consciously chosen to respond to a particular stimulus (Dickinson, 1995). This type of learning occurs when an agent is consciously aware of the consequences its actions bear, leading to an action (A) outcome (O) association. Contrary to the flexibility of goal-directed instrumental learning, research suggests that Pavlovian biases are often rigid and inflexible (Guitart-Masip, 2014). These biases are learned through classical conditioning mechanisms, which categorizes them under non-declarative and procedural learning. It has been argued that this type of learning arises automatically, without conscious awareness, and can persevere even when environmental context changes (Bouton et al., 2004). Moreover, recent research suggests that when instrumental learning and Pavlovian biases interact, executive cognitive functions decide which should take precedence by deducing which of the two processes might optimize the prediction of the reward in the given situation. This theory, called the 'arbitration theory', suggests that deployment of Pavlovian biases is favored when the reward is delivered under unpredictable situations (Moscarello et al., 2017). This means that when individuals are placed in situations where they have a low sense of control, they will rely on Pavlovian biases rather than instrumental learning (Dorfman et al., 2019). The arbitration theory is also consistent with learned helplessness research that focuses on the lack of or reduction in instrumental response when an individual senses a lack of control in escaping punishment (Alloy, 1982).

Our study consisted of two experiments, Experiment 1 and Experiment 2, both of which had three subparts: a) baseline (no reversal), b) reversal of instrumental learning and lastly, c) reversal of Pavlovian biases. Our goal with these two experiments was to assess how Pavlovian biases and goal-directed instrumental learning interact while making decisions as well as the flexibility of these two learning systems when contingencies are changed. For this study, we utilized a modified version of the go/no-go task where participants had to earn as many points as possible and correctly choose the action (pressing/not pressing) for the respective outcome (win/avoid loss) for each of the four differently colored robots – resulting in four trial types: Go to Win, Go to Avoid Loss, No Go to Win and No Go to Avoid Loss. For this study, we hypothesized that in comparison to goal-directed instrumental learning, Pavlovian biases would be slow to adjust to the reversal due to their reliance on non-declarative and procedural learning mechanisms. Experiment 1 & Experiment 2 had three overarching findings: presence and flexibility of Pavlovian biases, utilization of explicit information to verbalize valence for biases, and lastly, the persisting 'go' or action bias.

Key Findings

Influence of Pavlovian Bias & their Flexibility

Overpowering Biases

From the results of our analyses we can confirm that Pavlovian biases are present – often overpowering our goal-directed instrumental learning. For example, we saw a clear trend of Pavlovian biases overpowering goal-directed instrumental learning in conflicting contingencies (GAL & NGW) whereas when Pavlovian and instrumental learning were aligned, the average accuracy was distinctly higher (GW & NGAL). This general asymmetry in the results for the four trial types within the interaction of goal-directed instrumental learning with Pavlovian bias has been reported extensively for the Pavlovian go/no-go task (Raab & Hartley, 2020; Guitart-Masip et al., 2012; Hester, 2004; Wang et al., 2024). In 2012, Guitart-Masip and colleagues reported apparent interference of instrumental learning through Pavlovian biases (Guitart-Masip et al., 2012). They reported that the main reason for these inconsistencies in the learning was due to incongruency between valence and vigor of action, which was also evident in our results. Additionally, this trend is evident in animal research as well, for example, the classic study of Hershberger's chicks who were not able to learn that walking away from food would help them acquire it (Hershberger, 1986). Overarchingly, we frequently struggle to withhold response in the light of a predictive reward (NGW), while erroneously suppressing action when there is a necessity to act (GAL). This suggests that Pavlovian biases can significantly influence behavior, even in situations where instrumental learning is expected to dominate, for example 'No Go to Avoid Losing'.

Demonstration of Flexibility within Pavlovian Biases

The evidence for flexibility in Pavlovian biases comes from both Experiments 1c & 2c where we reversed Pavlovian associations (win became avoid loss and vice versa). When we reversed GAL to GW and NGW to NGAL performance improved (Figure 17 & 23). Although these biases have been known to exhibit fast and frugal actions – meaning they can be executed rapidly and are not cognitively demanding, there is still an ongoing debate about their flexibility properties (Boureau et al., 2015; Dayan et al., 2006). A recent study by Algermissen and colleagues proposed that goal-directed instrumental learning actively recruits Pavlovian biases by steering attention to the stimuli associated with these biases, which then allows the biases to affect behavior (Algermissen et al., 2023). This theory suggests that Pavlovian biases can adapt to situations especially when aligned with goal-directed instrumental learning. Specifically, if our goal-directed learning is selectively activating Pavlovian biases then we can find support for the claim that these biases can find synergy between automatic responses and more deliberate actions, providing evidence for our findings. With the consistent training that participants received after the reversal – where participants were able to understand the associations of the given trial type post reversal for 60 trials, comprising of 1 block – we were able to train participants on the new associations. As seen in Figure 17 and Figure 23, participants were able to rapidly regain their accuracy overtime. In a study conducted in 2021 researchers found that with sufficient training, Pavlovian biases can be diminished (Ereira et al., 2021). The authors found that in semantic spaces, specifically in verbal and motoric, these biases can shrink overtime with the modification of the task to be designed for optimal learning. Additionally, the change in the task in Experiment 2 to use deterministic relationships could have influenced participants' ability to override Pavlovian biases and express goal-directed actions. Research suggests that decreased sense of controllability or cognitive effort can lead to depression and anxiety – some of the maladaptive extensions of awry Pavlovian biases (Flemming et al., 2022). This could suggest that Pavlovian biases are considered as low control and low effort decision makers that can be regulated via cognitive effort and would be less likely to affect behavior in high perceived control environments characterized by deterministic relationships. Lastly, one possible mechanism that could account for the flexibility we observed is that the Pavlovian biases elicited in this task may actually be dependent on explicit learning – which was our second key finding.

Explicit vs Implicit Knowledge

After comparing participants post-task survey performance with their task performance, we found that the majority of the participants were able to verbalize the action and valence associated with a specific stimulus (Figure 33 & 34), and that their task performance was consistent with their post survey responses. Within the post-task survey, the mixed category outnumbered the other three categories as it had the most participants– meaning they were neither able to perfectly learn valence or action. These participants also showed suboptimal performance on the go/no-go task. The second most common category was the 'perfect' category where participants were able to not only answer the questions on 'actions' associated with a robot but also the associated 'valence' or outcome. 'Perfect' category participants showed an overall optimal performance on the go/no-go task aligning with their explicit knowledge. With further assessment of the post-task survey analysis, we found that participants who were categorized in the 'perfect valence only' category struggled to perform optimally on the task. Specifically, we noticed that participants were unable to correctly identify the association for incongruent trial types (GAL and NGW). The technique they employed to focus solely on the outcome a specific stimulus was producing without accounting for the actions needed to obtain that outcome became a hurdle when they encountered stimuli where actions did not match the outcome and were not reliable predictors of the corresponding action. For example, when winning was not associated with 'going' (NGW), or when avoiding loss was not a product of withholding response (GAL). This finding is worthy of exploration as to how the introduction of deterministic contingencies allowed participants to explicitly learn the association of a typical go/no-go task.

We posit that the simple nature of our paradigm – go no go task – as well as the coherent and straightforward instructions that were provided in the task for participants could have led to participants parsing out the associations of the trial types explicitly leading to a more conscious approach to the task (Nosek & Banaji, 2001). Similarly, we also provided participants with feedback throughout the experiment which might reinforce associations more easily, leading to a deliberate choice made by participants. Through the post-task survey, we found that participants gained explicit knowledge throughout the task, even for reversal of Pavlovian bias (refer to

Figure 32). These conditions may have supported explicit learning and flexible use of the Pavlovian biases, which is consistent with research finding that Pavlovian biases can be flexible depending on environmental contexts. In 2004, Bouton provided evidence concerning Pavlovian learning's flexible characteristics during extinction (Bouton, 2004). His study found that Pavlovian biases can be context-dependent, meaning that after extinction these biases might not appear in certain conditions/environments but emerge in new or different contexts. This offers an insight to the argument against complete rigidity of Pavlovian biases. Additionally, some researchers have also argued that perhaps, Pavlovian biases might not be fully rigid and store the ability to flexibly interact with instrumental learning especially in favorable contexts (Dayan et al., 2006). These theories might assist us in explaining the flexibility participants demonstrated during the reversal of Pavlovian biases (Experiment 1c & 2c). Recent research efforts provide evidence to support the role of conscious awareness within the working mechanisms of Pavlovian biases (Lovibons & Shanks, 2002; Madaboosi et al., 2012). In 2002, Lovibond and Shanks contradicted the traditional view by providing theoretical evidence that supported the role of declarative knowledge in Pavlovian biases. The authors found that these biases were not strictly rigid and automatic but had key aspects to them that allowed certain consciousness that employed awareness of stimuli and its respective contingencies (Lovibond & Shanks, 2002). Additionally, Madaboosi and colleagues discussed the role of awareness in fear conditioning (Madaboosi et al., 2012). With the focus on neurocircuitry involved in awareness, they found greater functional connectivity in the salience network during fear conditioning which provides evidence in favor of explicit and declarative awareness of stimuli contingencies in Pavlovian conditioning.

Persisting 'Go' Bias

Throughout this study we noticed a persistent 'go' bias as seen in the trend of highest average accuracy of 'go to win' in comparison to 'no go to avoid losing'. According to foundational research on the 'go' bias, individuals exhibit a noticeable preference for action over inaction in go/no-go tasks and in general when making decisions (Rosenblatt, 2003). The strength of this bias can be seen in reversal of GW to NGW in Pavlovian bias reversal in both Experiment 1 and 2. This phenomenon is termed 'action bias' in the current literature. Koster and colleagues stated that with the value of reward as an outcome, there is a pervasive intrinsic bias to act, indicating that choices involving action are preferred over inaction (Koster et al., 2015). Another work focuses on the cognitive ease of action (Albarracin et al., 2020). This study states that conducting an action is cognitively more straightforward than inhibiting a response introducing the concept of action dominance theory. This theory supports the idea that when individuals are placed in contexts where they have to balance between action and inaction continuously, to ease the cognitive load, they might focus on action solely which frequently might not be the most optimal choice.

Limitations

This study had some limitations that should be addressed in future research. First, there were concerns with the original go/no-go task that was utilized in this study. Due to participants inability to learn the associations correctly throughout the paradigm – 3 blocks of training – in Experiment 1, we had to modify aspects of the go/no-go task to optimize learning for participants. Our changes included an extra block of training, therefore a total of 4 blocks for each experiment (2a, 2b, 2c), which assisted in comprehending the associations better. Next, we altered the probability of the task associations from 80 – 20% to a completely deterministic task.

This eliminated the possibility for uncertainty for participants which increased their accuracy for Experiment 2. Additionally, we also condensed the time we allotted to each trial in a block by shortening the amount of time of the stimuli and feedback screens. This improved the momentum of our tasks and helped us add the aforementioned extra training block without increasing the time it took to complete our paradigm excessively. With these added modifications, participants' performance increased significantly and led to robust results, however, due to all the additional changes we could have unintentionally introduced variability and limited the generalizability of our results. Our study also had a limited demographics population which was recruited from Colorado State University. Due to the restricted recruitment, we could have unintentionally overlooked how various subgroups of populations belonging to distinct backgrounds, socioeconomic status, race and age may affect the results of our findings. Lastly, we would like to acknowledge that in Experiment 1, we ran our three studies (1a, 1b, and 1c) sequentially instead of simultaneously. This could have introduced variability in our data albeit we did adjust for this in our Experiment 2 by randomly assigning participants to conditions to ensure robustness of our results.

Study Implications

We can conjure several implications from the results of our study. Firstly, this study adds to the literature on the interplay between goal-directed instrumental learning and Pavlovian bias. In this study we uncovered the overpowering tendencies of our learning biases when a decision requires a goal-directed action and how they can prevail over our more conscious actions as shown by the compromised performance on trial types of GAL and NGW. Secondly, we found that Pavlovian biases exhibit flexibility in certain contexts. This was evident during the reversal of Pavlovian biases (Experiment 1c and 2c) where a subgroup of participants were able to verbalize the valence associated to a specific robot and explicitly modify their associations post reversal. This contributes to the foundational research that argues that Pavlovian biases can exhibit flexibility based on the contexts that they are elicited in. Therefore, our study extends to a more cohesive discussion on how explicit and implicit information may work during certain contexts and decision-making circumstances. We also found plethora of evidence that suggested the effect of individual differences on performance that were demonstrated by participants in our go/no-go paradigm. This idiosyncratic variability in participants performance in the task suggested how various cognitive functions such as attention, innate cognitive flexibility and heuristics can play a crucial role when forming a decision – leading to broader implications.

Rigidity in Pavlovian biases is known to be associated with various decision making disorders such as Substance Use Disorders (SUDs), Obsessive Compulsive Disorder (OCD), anxiety, depression, gambling as well as eating disorders. Pavlovian biases can sensitize individuals to stimuli that can disrupt their conscious and intentional actions when making a decision which causes them to rely on their readily available automatic biases. This study contributes to foundational and theoretical framework to understand how learning biases strengthen these disorders and continuously maintain them to disrupt executive functioning. With this study we hope to contribute to the limited understanding of Pavlovian biases rigidity, especially during their interaction with our conscious goal-directed actions. Additionally, this study elucidates the significance of cognitive flexibility and its importance in preventing cognitive disorders stemming from dysregulated biases. We hope to emphasize the importance of early interventions such as meditation or cognitive therapy to address cognitive disorders that emerge due to erroneous decision making.

Future Directions

We believe that the implications of the results found in both Experiment 1 and Experiment 2 can be employed in a variety of settings. Firstly, this study informs us of the intricacies of decision making and human learning – with this study we examined how learning biases may affect our goal-directedness and impede on our ability to make optimal decisions. With this foundation of understanding we can also tackle how Pavlovian bias engages in flexibility and adaptability in certain situations. Since in our reversal of Pavlovian biases we noticed that participants were able to rapidly adjust to the reversal, it informed us of two things which could be explored further: 1) the reliability of the go/no-go task to test the rigidity of Pavlovian biases 2) more evidence supporting the flexibility demonstrated by Pavlovian biases in certain environments such as low or high control situations. Additionally, we could also explore how the persistent go biases that participants relied on throughout this study are facilitated by Pavlovian biases and their interaction. In the future, we hope to expand the understanding from this study to a variety of clinical populations dealing with disorders that stem from maladaptive extensions of Pavlovian biases such as Substance Use Disorders (SUDs), addictions, Obsessive Compulsive Disorder (OCD) as well as anxiety and depression. Lastly, we aim to utilize functional magnetic resonance imaging (fMRI) to explore the neural circuitry involved in these learning biases and overlapping decision making neural systems.

REFERENCES

Algermissen, & Hanneke. (2023). Goal-directed recruitment of Pavlovian biases through selective visual attention. Journal of Experimental Psychology General, 152(10), 2941–2956. https://doi.org/10.1037/xge0001425

Albarracín, D., & Dai, W. (2024). The impact of the environment on behavior. *Advances in Experimental Social Psychology*, *69*. https://doi.org/10.1016/bs.aesp.2023.12.001

Alloy, L. B. (1982). The role of perceptions and attributions for response-outcome noncontingency in learned helplessness: A commentary and discussion. *Journal of Personality*, *50*(4), 443–479. https://doi.org/10.1111/j.1467-6494.1982.tb00229.x

Ashby, F. G., & O'Brien, J. B. (2005). Category learning and multiple memory systems. Trends in Cognitive Sciences, 9(2), 83–89. https://doi.org/10.1016/j.tics.2004.12.003

Avena, N. M., & Bocarsly, M. E. (2011). Dysregulation of brain reward systems in eating disorders: Neurochemical information from animal models of binge eating, bulimia nervosa, and anorexia nervosa. *Neuropharmacology*, *63*(1), 87–96. https://doi.org/10.1016/j.neuropharm.2011.11.010

Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. Neuropharmacology, 37(4-5), 407-419. https://doi.org/10.1016/S0028-3908(98)00033-1

Balleine, B. W., & O'Doherty, J. P. (2010). Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology, 35(1), 48-69. https://doi.org/10.1038/npp.2009.131

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? Brain Research Reviews, 28(3), 309-369. https://doi.org/10.1016/S0165-0173(98)00019-8

Bitterman, M. E. (2006). Classical conditioning since Pavlov. Integrative and Comparative Biology, 46(3), 240-250. https://doi.org/10.1037/1089-2680.10.4.365

Bray, S., Rangel, A., Shinsuke Shimojo, Balleine, B., & O'Doherty, J. P. (2008). The Neural Mechanisms Underlying the Influence of Pavlovian Cues on Human Decision Making. *Journal of Neuroscience*, *28*(22), 5861–5866. https://doi.org/10.1523/jneurosci.0897-08.2008

Bouras, N. N., Mack, N. R., & Gao, W.-J. (2023). Prefrontal modulation of anxiety through a lens of noradrenergic signaling. Frontiers in Systems Neuroscience, 17. https://doi.org/10.3389/fnsys.2023.1173326

Boureau, Y-Lan., Sokol-Hessner, P., & Daw, N. D. (2015). Deciding How To Decide: Self-Control and Meta-Decision Making. *Trends in Cognitive Sciences*, *19*(11), 700–710. https://doi.org/10.1016/j.tics.2015.08.013

Bouton, M. E. (2004). Context and behavioral processes in extinction. Learning & Memory, 11(5), 485-494. https://doi.org/10.1101/lm.78804

Bouton, M. E., & Peck, C. A. (1992). Spontaneous recovery in cross-motivational transfer (counterconditioning). Animal Learning & Behavior, 20(4), 313–321. https://doi.org/10.3758/bf03197954

Cartoni, E., Balleine, B., & Baldassarre, G. (2016). Appetitive Pavlovian-instrumental Transfer: A review. Neuroscience & Biobehavioral Reviews, 71, 829–848. https://doi.org/10.1016/j.neubiorev.2016.09.020

Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M. J. (2013). Frontal Theta Overrides Pavlovian Learning Biases. Journal of Neuroscience, 33(19), 8541–8548. https://doi.org/10.1523/jneurosci.5754-12.2013

Chen, A. C., & Etkin, A. (2013). Hippocampal Network Connectivity and Activation Differentiates Post-Traumatic Stress Disorder From Generalized Anxiety Disorder. Neuropsychopharmacology, 38(10), 1889–1898. https://doi.org/10.1038/npp.2013.122

Chen, Q., Liu, M., Wen, R., Xu, C., Wei, Z., Zhang, W., Seger, C. A., & Peng, Z. (2022). Dysfunction of goal-directed control in patients with depression and nonsuicidal self-injury. Brain and Behavior, 12(6). https://doi.org/10.1002/brb3.2607

Childress. (2016). Cue reactivity and cue reactivity interventions in drug dependence. NIDA Research Monograph, 137. https://pubmed.ncbi.nlm.nih.gov/8289929/

Clark, J. J., Hollon, N. G., & Phillips, P. E. M. (2012). Pavlovian valuation systems in learning and decision making. Current Opinion in Neurobiology, 22(6), 1054-1061. https://doi.org/10.1016/j.conb.2012.06.004

Crockett, M. J., Clark, L., Golnaz Tabibnia, Lieberman, M. D., & Robbins, T. W. (2008). Serotonin Modulates Behavioral Reactions to Unfairness. Science, 320(5884), 1739–1739. https://doi.org/10.1126/science.1155577

Dalton, G. L., Phillips, A. G., & Floresco, S. B. (2014). Preferential Involvement by Nucleus Accumbens Shell in Mediating Probabilistic Learning and Reversal Shifts. *The Journal of Neuroscience*, *34*(13), 4618–4626. https://doi.org/10.1523/jneurosci.5058-13.2014

Dajani, D. R., & Uddin, L. Q. (2015). Demystifying cognitive flexibility: Implications for clinical and developmental neuroscience. *Trends in Neurosciences*, *38*(9), 571–578. https://doi.org/10.1016/j.tins.2015.07.003

Dayan, P., & Seymour, B. (2009). Values and Actions in Aversion. Elsevier EBooks, 175–191. https://doi.org/10.1016/b978-0-12-374176-9.00012-9

Dere, E., & Armin Zlomuzica. (2014). Far Beyond Declarative and Non-Declarative Memories. Frontiers in Behavioral Neuroscience, 8. https://doi.org/10.3389/fnbeh.2014.00173

Dickinson, A. (1985). Actions and habits: the development of behavioural autonomy. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 308(1135), 67–78. https://doi.org/10.1098/rstb.1985.0010

D'Souza, D, & Amrutha Avati. (2021). Memory and Learning: Basic Concepts. Springer EBooks, 227–240. https://doi.org/10.1007/978-3-030-72400-9_11

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. Neuron, 80(2), 312-325. https://doi.org/10.1016/j.neuron.2013.09.007

Domjan, M. (2005). Pavlovian Conditioning: A Functional Perspective. Annual Review of Psychology, 56(1), 179–206. https://doi.org/10.1146/annurev.psych.55.090902.141409

Ebrahimi, C., Garbusow, M., Sebold, M., Chen, K., Smolka, M. N., Huys, Q. J. M., Zimmermann, U. S., Schlagenhauf, F., & Heinz, A. (2023). Elevated Amygdala Responses During De Novo Pavlovian Conditioning in Alcohol Use Disorder Are Associated With Pavlovian-to-Instrumental Transfer and Relapse Latency. *Biological Psychiatry Global Open Science*, *3*(4), 803–813. https://doi.org/10.1016/j.bpsgos.2023.02.003

Dorfman, H. M., & Gershman, S. J. (2019). Controllability governs the balance between Pavlovian and instrumental action selection. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-13737-7

Eichenbaum, H. (2000). A cortical-hippocampal system for declarative memory. Nature Reviews Neuroscience, 1(1), 41-50. https://doi.org/10.1038/35036213

Ereira, S., Pujol, M., Guitart-Masip, M., Dolan, R., & Zeb Kurth-Nelson. (2021). Overcoming Pavlovian bias in semantic space. Scientific Reports, 11(1). https://doi.org/10.1038/s41598-021-82889-8

Eryilmaz, Anais Rodriguez-Thompson, Tanner, A. S., Giegold, M., Huntington, F. C., & Roffman, J. L. (2017). Neural determinants of human goal-directed vs. habitual action control and their relation to trait motivation. Scientific Reports, 7(1). https://doi.org/10.1038/s41598-017-06284-y

Everitt, B. J., & Robbins, T. W. (2005). Neural systems of reinforcement for drug addiction: From actions to habits to compulsion. Nature Neuroscience, 8(11), 1481-1489. https://doi.org/10.1038/nn1579

Fabio, R. A., Tindara Caprì, & Romano, M. (2019). From Controlled to Automatic Processes and Back Again: The Role of Contextual Features. Europe's Journal of Psychology, 15(4), 773–788. https://doi.org/10.5964/ejop.v15i4.1746

Farashahi, S., Donahue, C. H., Hayden, B. Y., Lee, D., & Soltani, A. (2019). Flexible combination of reward information across primates. Nature Human Behaviour, 3(11), 1215–1224. https://doi.org/10.1038/s41562-019-0714-3

Fischer, S. (2021). The hypothalamus in anxiety disorders. *Handbook of Clinical Neurology*, 149–160. https://doi.org/10.1016/b978-0-12-820107-7.00009-4

Fleming, H. A., Feng, G., Rutledge, R., Roiser, J. P., & Robinson, O. J. (2023). Training successfully reduces the strength of Pavlovian biases. https://doi.org/10.31234/osf.io/h3297

Foerde, K., & Shohamy, D. (2011). The role of the basal ganglia in learning and memory: Insight from Parkinson's disease. Neurobiology of Learning and Memory, 96(4), 624–636. https://doi.org/10.1016/j.nlm.2011.08.006

Fordyce, W. E., & Company, M. C. V. (1977). Behavioral methods for chronic pain and illness. *Pain*, *3*(3), 291–292. https://doi.org/10.1016/0304-3959(77)90029-x

Garland, E. L., Franken, A., & Howard, M. O. (2011). Cue-elicited heart rate variability and attentional bias predict alcohol relapse following treatment. Psychopharmacology, 222(1), 17–26. https://doi.org/10.1007/s00213-011-2618-4

Garbusow, M., Ebrahimi, C., Riemerschmid, C., Daldrup, L., Rothkirch, M., Chen, K., Chen, H., Belanger, M. J., Hentschel, A., Smolka, M. N., Heinz, A., Maximilan Pilhatsch, & Rapp, M. A. (2022). Pavlovian-to-Instrumental Transfer across Mental Disorders: A Review. Neuropsychobiology, 81(5), 418–437. https://doi.org/10.1159/000525579

Gillan, C. M., Papmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T. W., & Sanne de Wit. (2011). Disruption in the Balance Between Goal-Directed Behavior and Habit Learning in Obsessive-Compulsive Disorder. American Journal of Psychiatry, 168(7), 718–726. https://doi.org/10.1176/appi.ajp.2011.10071062

Glasner, S., & Drazdowski, T. K. (2018). Evidence-based Behavioral Treatments for Substance Use Disorders. Elsevier EBooks, 157–166. https://doi.org/10.1016/b978-0-323-54856-4.00010-9

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. Annual Review of Neuroscience, 31, 359-387. https://doi.org/10.1146/annurev.neuro.29.051605.112851

Gremel, C. M., & Costa, R. M. (2013). Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. Nature Communications, 4(1). https://doi.org/10.1038/ncomms3264

Gruner, P., Anticevic, A., Lee, D., & Pittenger, C. (2015). Arbitration between Action Strategies in Obsessive-Compulsive Disorder. The Neuroscientist, 22(2), 188–198. https://doi.org/10.1177/1073858414568317

Gold, J. I., & Shadlen, M. N. (2007). The Neural Basis of Decision Making. Annual Review of Neuroscience, 30(1), 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Guitart-Masip, M., Huys, Q. J. M., Lluis Fuentemilla, Dayan, P., Emrah Duzel, & Dolan, R. J. (2012). Go and no-go learning in reward and punishment: Interactions between affect and effect. NeuroImage, 62(1), 154–166. https://doi.org/10.1016/j.neuroimage.2012.04.024

Guitart-Masip, M., Emrah Duzel, Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. Trends in Cognitive Sciences, 18(4), 194–202. https://doi.org/10.1016/j.tics.2014.01.003

Hansson, A. C., Koopmann, A., Uhrig, S., Bühler, S., Esi Domi, Kiessling, E., Ciccocioppo, R., Froemke, R. C., Grinevich, V., Kiefer, F., Sommer, W. H., Vollstädt-Klein, S., & Rainer Spanagel. (2017). Oxytocin Reduces Alcohol Cue-Reactivity in Alcohol-Dependent Rats and Humans. Neuropsychopharmacology, 43(6), 1235–1246. https://doi.org/10.1038/npp.2017.257

Harris, C., Aguirre, C., Kolli, S., Das, K., Izquierdo, A., & Soltani, A. (2021). Unique features of stimulus-based probabilistic reversal learning. Behavioral Neuroscience, 135(4), 550–570. https://doi.org/10.1037/bne0000474

Harvie, D. S., Moseley, G. L., Hillier, S. L., & Meulders, A. (2017). Classical Conditioning Differences Associated With Chronic Pain: A Systematic Review. *The Journal of Pain*, *18*(8), 889–898. https://doi.org/10.1016/j.jpain.2017.02.430

Helmstetter, F. J., & Bellgowan, P. S. (1994). Effects of muscimol applied to the amygdala on Pavlovian fear conditioning in rats. Behavioral Neuroscience, 108(5), 1005-1009. https://doi.org/10.1037/0735-7044.108.5.1005

Hester, R. (2004). Individual Differences in Error Processing: A Review and Reanalysis of Three Event-related fMRI Studies Using the GO/NOGO Task. *Cerebral Cortex*, *14*(9), 986–994. https://doi.org/10.1093/cercor/bhh059

Hershberger, W. A. (1986). An approach through the looking-glass. Animal Learning & Behavior, 14(4), 443–451. https://doi.org/10.3758/bf03200092

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. Memory & Cognition, 10(4), 389-395. https://doi.org/10.3758/BF03202431

Holmes, N. M., Marchand, A. R., & Coutureau, E. (2010). Pavlovian to instrumental transfer: A neurobehavioural perspective. Neuroscience & Biobehavioral Reviews, 34(8), 1277-1295. https://doi.org/10.1016/j.neubiorev.2010.03.007

Huys, Guitart-Masip, M., Dolan, R. J., & Dayan, P. (2015). Decision-Theoretic Psychiatry. Clinical Psychological Science, 3(3), 400–421. https://doi.org/10.1177/2167702614562040

Ishikawa, A., F. Ambroggi, Nicola, S. M., & Fields, H. L. (2008). Contributions of the amygdala and medial prefrontal cortex to incentive cue responding. Neuroscience, 155(3), 573–584. https://doi.org/10.1016/j.neuroscience.2008.06.037

Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux. Kahneman, D. (2011). Thinking, fast and slow. *APA PsycNet*. (2025). Apa.org. https://psycnet.apa.org/record/2011-26535-000

Kaye, W. H., Fudge, J. L., & Paulus, M. (2009). New insights into symptoms and neurocircuit function of anorexia nervosa. Nature Reviews. Neuroscience, 10(8), 573–584. https://doi.org/10.1038/nrn2682

Kim, D., Park, G. Y., O′Doherty, J. P., & Lee, S. W. (2019). Task complexity interacts with state-space uncertainty in the arbitration between model-based and model-free learning. Nature Communications, 10(1). https://doi.org/10.1038/s41467-019-13632-1

Kiefer, M. (2012). Executive control over unconscious cognition: attentional sensitization of unconscious information processing. Frontiers in Human Neuroscience, 6. https://doi.org/10.3389/fnhum.2012.00061

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. Science, 273(5280), 1399-1402. https://doi.org/10.1126/science.273.5280.1399

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. Science, 303(5660), 1023-1026. https://doi.org/10.1126/science.1089910

Lange, F., Seer, C., & Kopp, B. (2017). Cognitive flexibility in neurological disorders: Cognitive components and event-related potentials. *Neuroscience & Biobehavioral Reviews*, *83*, 496–507. https://doi.org/10.1016/j.neubiorev.2017.09.011

LeDoux, J. E. (2000). Emotion circuits in the brain. Annual Review of Neuroscience, 23, 155-184. https://doi.org/10.1146/annurev.neuro.23.1.155

Lee, S., Shinsuke Shimojo, & John P. O'Doherty. (2014). Neural Computations Underlying Arbitration between Model-Based and Model-free Learning. Neuron, 81(3), 687–699. https://doi.org/10.1016/j.neuron.2013.11.028

Lovibond, P. F., & Shanks, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. Journal of Experimental Psychology: Animal Behavior Processes, 28(1), 3-26. https://doi.org/10.1037/0097-7403.28.1.3

Madaboosi, S, Grasser, L. R., Chowdury, A & Javanbakht, A. (2021). Neurocircuitry of Contingency Awareness in Pavlovian Fear Conditioning. Cognitive Affective & Behavioral Neuroscience, 21(5), 1039–1053. https://doi.org/10.3758/s13415-021-00909-6

Martin, E. I., Ressler, K. J., Binder, E., & Nemeroff, C. B. (2009). The Neurobiology of Anxiety Disorders: Brain Imaging, Genetics, and Psychoneuroendocrinology. Psychiatric Clinics of North America, 32(3), 549–575. https://doi.org/10.1016/j.psc.2009.05.004

Mayberg, H. S. (1997). Limbic-cortical dysregulation: a proposed model of depression. Journal of Neuropsychiatry, 9(3), 471–481. https://doi.org/10.1176/jnp.9.3.471

Menzies, L., Achard, S., Chamberlain, S. R., Fineberg, N., Chen, C.-H., Campo, N. del, Sahakian, B. J., Robbins, T. W., & Bullmore, E. (2007). Neurocognitive endophenotypes of obsessive-compulsive disorder. Brain, 130(12), 3223–3236. https://doi.org/10.1093/brain/awm205

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. Annual Review of Neuroscience, 24, 167-202. https://doi.org/10.1146/annurev.neuro.24.1.167

Minto, N., Stella, M., & McIlvane, W. J. (2014). Discrimination and Reversal Learning by Toddlers Aged 15-23 Months. The Psychological Record, 65(1), 41–47. https://doi.org/10.1007/s40732-014-0084-1

Mkrtchian, A., Aylward, J., Dayan, P., Roiser, J. P., & Robinson, O. J. (2017). Modeling Avoidance in Mood and Anxiety Disorders Using Reinforcement Learning. Biological Psychiatry, 82(7), 532–539. https://doi.org/10.1016/j.biopsych.2017.01.017

Muller, J., Corodimas, K. P., Fridel, Z., & LeDoux, J. E. (1997). Functional inactivation of the lateral and basal nuclei of the amygdala by muscimol infusion prevents fear conditioning to an explicit conditioned stimulus and to contextual stimuli. Behavioral Neuroscience, 111(4), 683–691. https://doi.org/10.1037//0735-7044.111.4.683

Modecki, K. L., Ryan, K. M., & Waters, A. M. (2022). Fear learning and extinction predicts anxiety in daily life: a study of Pavlovian conditioning and ecological momentary assessment. *Psychological Medicine*, *53*(11), 5301–5311. https://doi.org/10.1017/s0033291722002379

Moscarello, J. M., & Hartley, C. A. (2017). Agency and the Calibration of Motivated Behavior. *Trends in Cognitive Sciences*, *21*(10), 725–735. https://doi.org/10.1016/j.tics.2017.06.008

Moutoussis, M., Bullmore, E. T., Goodyer, I. M., Fonagy, P., Jones, P. B., Dolan, R. J., & Dayan, P. (2018). Change, stability, and instability in the Pavlovian guidance of behaviour from adolescence to young adulthood. PLoS Computational Biology, 14(12), e1006679–e1006679. https://doi.org/10.1371/journal.pcbi.1006679

Myers, K. M., & Davis, M. (2007). Mechanisms of fear extinction. Molecular Psychiatry, 12(2), 120-150. https://doi.org/10.1038/sj.mp.4001939

Nadel, L. (1992). Multiple Memory Systems: What and Why. Journal of Cognitive Neuroscience, 4(3), 179–188. https://doi.org/10.1162/jocn.1992.4.3.179

Nadler, N., Delgado, M. R., & Delamater, A. R. (2011). Pavlovian to instrumental transfer of control in a human learning task. Emotion, 11(5), 1112–1123. https://doi.org/10.1037/a0022760

Nestler, E. J., & Carlezon, W. A. (2006). The Mesolimbic Dopamine Reward Circuit in Depression. Biological Psychiatry, 59(12), 1151–1159. https://doi.org/10.1016/j.biopsych.2005.09.018

Vafaie, N., & Kober, H. (2022). Association of Drug Cues and Craving With Drug Use and Relapse. *JAMA Psychiatry*, *79*(7), 641. https://doi.org/10.1001/jamapsychiatry.2022.1240

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2006). Tonic dopamine: opportunity costs and the control of response vigor. Psychopharmacology, 191(3), 507–520. https://doi.org/10.1007/s00213-006-0502-4

Nosek, B. A., & Banaji, M. R. (2001). The Go/No-Go Association Task. *Social Cognition*, *19*(6), 625–666. https://doi.org/10.1521/soco.19.6.625.20886

O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: Insights from neuroimaging. Current Opinion in Neurobiology, 14(6), 769-776. https://doi.org/10.1016/j.conb.2004.10.016

Pakaprot. (2025). Eyeblink Conditioning As Models of Declarative and Non-declarative Memories. *Journal of Physiological and Biomedical Sciences*, *24*(2), 1–12. https://li01.tci-thaijo.org/index.php/j-pbs/article/view/251014

Park, H., Doh, H., Lee, E., Park, H., & Ahn, W.-Y. (2023). The neurocognitive role of working memory load when Pavlovian motivational control affects instrumental learning. PLoS Computational Biology, 19(12), e1011692–e1011692. https://doi.org/10.1371/journal.pcbi.1011692

Peng, Z., He, L., Wen, R., Verguts, T., Seger, C. A., & Chen, Q. (2022). Obsessive-compulsive disorder is characterized by decreased Pavlovian influence on instrumental behavior. PLOS Computational Biology/PLoS Computational Biology, 18(10), e1009945–e1009945. https://doi.org/10.1371/journal.pcbi.1009945

Raab, H. A., & Hartley, C. A. (2020). Adolescents exhibit reduced Pavlovian biases on instrumental learning. Scientific Reports, 10(1). https://doi.org/10.1038/s41598-020-72628-w

Raposo-Lima, C., & Morgado, P. (2020). The Role of Stress in Obsessive-Compulsive Disorder: A Narrative Review. Harvard Review of Psychiatry, 28(6), 356–370. https://doi.org/10.1097/hrp.0000000000000274 Reber, P. J., Knowlton, B. J., & Squire, L. R. (1996). Dissociable properties of memory systems: Differences in the flexibility of declarative and nondeclarative knowledge. Behavioral Neuroscience, 110(5), 861–871. https://doi.org/10.1037//0735-7044.110.5.861

Rehman, I., Navid Mahabadi, Sanvictores, T., & Rehman, C. I. (2024, September 5). Classical Conditioning. Nih.gov; StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK470326/

Rescorla, R. A. (1988). Behavioral studies of Pavlovian conditioning. Annual Review of Neuroscience, 11(1), 329-352. https://doi.org/10.1146/annurev.ne.11.030188.001553

Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. Psychological Review, 74(3), 151-182. https://doi.org/10.1037/h0024475

Remijnse, P. L., Marjan M. A. Nielen, Anton, Cath, D. C., Oppen, P. van, Harry, & Veltman, D. J. (2006). Reduced Orbitofrontal-Striatal Activity on a Reversal Learning Task in Obsessive-Compulsive Disorder. *Archives of General Psychiatry*, *63*(11), 1225–1225. https://doi.org/10.1001/archpsyc.63.11.1225

Rolf-Detlef, T. (2018). The International Association for the Study of Pain definition of pain: as valid in 2018 as in 1979, but in need of regularly updated footnotes. *PAIN Reports*, *3*(2), e643. https://doi.org/10.1097/pr9.0000000000000643

Rosenblatt, Bonnie L. (2003). Preferences for action or inaction in decision-making.. University of Illinois Chicago. Thesis. https://hdl.handle.net/10027/18502

Roy, M. H., & Dugal, S. S. (1998). Developing trust: The importance of cognitive flexibility and co-operative contexts. Management Decision, 36(9), 561
567. https://doi.org/10.1108/00251749810239441

Rudnik, A., Piotrowicz, G., Małgorzata Anna Basińska, & Vahid Rashedi. (2019). The importance of cognitive flexibility and flexibility in coping with stress for the quality of life in inflammatory bowel disease patients during biological therapy. A preliminary report. *Gastroenterology Review*, *14*(2), 121–128. https://doi.org/10.5114/pg.2018.81081

Sayette, M. A., Kirchner, T. R., Moreland, R. L., Levine, J. M., & Travis, T. (2004). Effects of Alcohol on Risk-Seeking Behavior: A Group-Level Analysis. Psychology of Addictive Behaviors, 18(2), 190–193. https://doi.org/10.1037/0893-164x.18.2.190

Schultz, W. (1998). Predictive Reward Signal of Dopamine Neurons. Journal of Neurophysiology, 80(1), 1–27. https://doi.org/10.1152/jn.1998.80.1.1

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. Annual Review of Psychology, 57(1), 87-115. https://doi.org/10.1146/annurev.psych.56.091103.070229

Sebold, M., Lorenz Deserno, Nebe, S., Schad, D. J., Garbusow, M., Hägele, C., Keller, J., Jünger, E., Kathmann, N., Smolka, M., Rapp, M. A., Schlagenhauf, F., Heinz, A., & Quentin J.M. Huys. (2014). Model-Based and Model-Free Decisions in Alcohol Dependence. Neuropsychobiology, 70(2), 122–131. https://doi.org/10.1159/000362840

Schoenbaum, G., Roesch, M. R., Stalnaker, T. A., & Takahashi, Y. K. (2009). A new perspective on the role of the orbitofrontal cortex in adaptive behaviour. Nature Reviews Neuroscience, 10(12), 885-892. https://doi.org/10.1038/nrn2753

Shanks, D. R. (1993). Human instrumental learning: A critical review of data and theory. British Journal of Psychology, 84(3), 319–354. https://doi.org/10.1111/j.2044-8295.1993.tb02486.x

Sherry, D. F., & Schacter, D. L. (1987). The evolution of multiple memory systems. Psychological Review, 94(4), 439–454. https://doi.org/10.1037/0033-295X.94.4.439

Skinner, B. F. (1953). Some contributions of an experimental analysis of behavior to psychology as a whole. American Psychologist, 8(2), 69–78. https://doi.org/10.1037/h0054118

Steding, J., Boehm, I., King, J. A., Geisler, D., Ritschel, F., Seidel, M., Doose, A., Jaite, C., Roessner, V., Smolka, M. N., & Ehrlich, S. (2019). Goal-directed vs. habitual instrumental behavior during reward processing in anorexia nervosa: an fMRI study. Scientific Reports, 9(1). https://doi.org/10.1038/s41598-019-49884-6

Stein, D. J. (2002). Obsessive-compulsive disorder. The Lancet, 360(9330), 397–405. https://doi.org/10.1016/s0140-6736(02)09620-4

Shin, L. M., & Liberzon, I. (2010). The neurocircuitry of fear, stress, and anxiety disorders. Neuropsychopharmacology, 35(1), 169-191. https://doi.org/10.1038/npp.2009.83

Sander C.J. Verfaillie, Stella, Vriend, C., Remijnse, P. L., Veltman, D. J., & van. (2016). The course of the neural correlates of reversal learning in obsessive–compulsive disorder and major depression: A naturalistic follow-up fMRI study. Journal of Obsessive-Compulsive and Related Disorders, 9, 51–58. https://doi.org/10.1016/j.jocrd.2016.02.004

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? Behavioral and Brain Sciences, 23(5), 645–665. https://doi.org/10.1017/S0140525X00003435

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. Neurobiology of Learning and Memory, 82(3), 171-177. https://doi.org/10.1016/j.nlm.2004.06.005

Squire, L. R. (2009). The Legacy of Patient H.M. for Neuroscience. Neuron, 61(1), 6–9. https://doi.org/10.1016/j.neuron.2008.12.023

Takahashi, T. T., Vengeliene, V., Enkel, T., Reithofer, S., & Rainer Spanagel. (2019). Pavlovian to Instrumental Transfer Responses Do Not Correlate With Addiction-Like Behavior in Rats. Frontiers in Behavioral Neuroscience, 13. https://doi.org/10.3389/fnbeh.2019.00129

Toner, J., Montero, B. G., & Moran, A. (2015). The Perils of Automaticity. *Review of General Psychology*, *19*(4), 431–442. https://doi.org/10.1037/gpr0000054

Tulving, E. (1985). How many memory systems are there? American Psychologist, 40(4), 385-398. https://doi.org/10.1037/0003-066X.40.4.385

Vafaie, N., & Kober, H. (2022). Association of Drug Cues and Craving With Drug Use and Relapse. JAMA Psychiatry, 79(7), 641. https://doi.org/10.1001/jamapsychiatry.2022.1240

Voon, Derbyshire, K., C Rück, Irvine, M. A., Y Worbe, J Enander, Schreiber, L. R. N., Gillan, C., Fineberg, N. A., Sahakian, B. J., Robbins, T. W., Harrison, N. A., Wood, J., Daw, N. D., Dayan, P., Grant, J. E., & Bullmore, E. T. (2014). Disorders of compulsivity: a common bias towards learning habits. Molecular Psychiatry, 20(3), 345–352. https://doi.org/10.1038/mp.2014.44

Waldron, E. M., & Ashby, F. G. (2001). The effects of concurrent task interference on category learning: Evidence for multiple category learning systems. Psychonomic Bulletin & Review, 8(1), 168–176. https://doi.org/10.3758/bf03196154

Watson, P., & Mahlberg, J. (2023). Mechanisms underlying performance in a cued go/no-go Pavlovian-to-instrumental transfer paradigm. Behavioural Brain Research, 446, 114413–114413. https://doi.org/10.1016/j.bbr.2023.114413

Wang, Z., Liu, X., & Li, X. (2023). EXPRESS: Unraveling the Dynamics of Response Force: Investigating Motor Restraint and Motor Cancellation Through Go/No-Go and Stop-Signal Tasks. *Quarterly Journal of Experimental Psychology*, *77*(11), 2199–2213. https://doi.org/10.1177/17470218231219867

Weiss, E. O., Kruppa, J. A., Fink, G. R., Beate Herpertz-Dahlmann, Konrad, K., & Schulte-Rüther, M. (2021). Developmental Differences in Probabilistic Reversal Learning: A Computational Modeling Approach. Frontiers in Neuroscience, 14. https://doi.org/10.3389/fnins.2020.536596

Wiggert, N., Wilhelm, F. H., Boger, S., Georgii, C., Wolfgang Klimesch, & Jens Blechert. (2016). Social Pavlovian conditioning: Short- and long-term effects and the role of anxiety and depressive symptoms. *Social Cognitive and Affective Neuroscience*, *12*(2), 329–339. https://doi.org/10.1093/scan/nsw128

Wimmer, G. E., & Shohamy, D. (2012). Preference by Association: How Memory Mechanisms in the Hippocampus Bias Decisions. Science, 338(6104), 270–273. https://doi.org/10.1126/science.1223252

Whitton, A. E., Treadway, M. T., & Pizzagalli, D. A. (2015). Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. Current Opinion in Psychiatry, 28(1), 7–12. https://doi.org/10.1097/yco.0000000000000122

Yin, H. H., & Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. Nature Reviews. Neuroscience, 7(6), 464–476. https://doi.org/10.1038/nrn1919

Zech, H. G., Rotteveel, M., Dijk, van, & Dillen, van. (2020). A mobile approach-avoidance task. Behavior Research Methods, 52(5), 2085–2097. https://doi.org/10.3758/s13428-020-01379-3

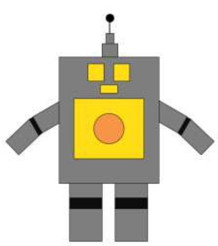
Zhang, D., Lee, E. K. P., Mak, E. C. W., Ho, C. Y., & Wong, S. Y. S. (2021). Mindfulness-based interventions: an overall review. *British Medical Bulletin*, *138*(1), 41–57. https://doi.org/10.1093/bmb/ldab005

APPENDIX A

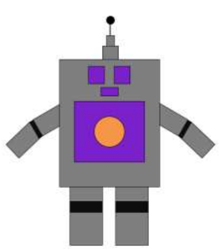
Baseline Post-Task Survey

1. Please describe the strategies you used, if any, to learn the correct response for ed to answer each robot

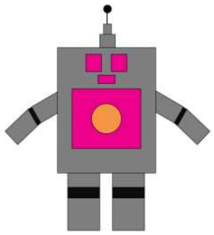
2. For this question, please disregard what the best response was (whether you think you should press the space bar, or refrain from pressing the space) and just let us know whether the robot was usually associated with winning or losing by circling the word "win" or "loss" If you think that the robot was associated with both winning and losing at times, you can circle the word "both". If you don't remember, circle "Don't know".



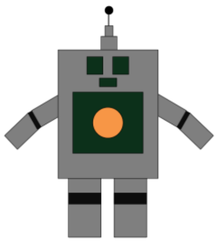
Win Loss Both Don't know



Win Loss Both Don't know



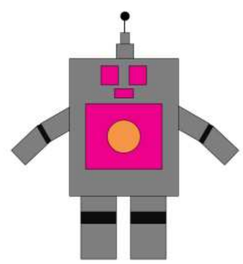
Win | Loss | Both | Don't know



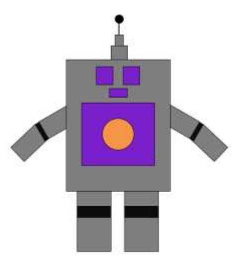
Win | Loss | Both | Don't know

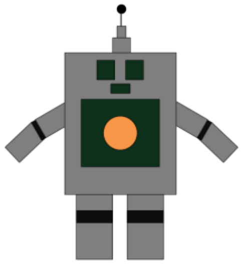
3. For some robots, the best response was to press the space bar (to win, or to avoid losing), whereas for other robots the best response was to not press the space bar. For this question, please disregard whether the robot was associated with a win or loss and just indicate whether the best response is to press the space bar or not press for this robot. If you think the two responses are equally good, you can choose the word 'Either'. Circle the best option.

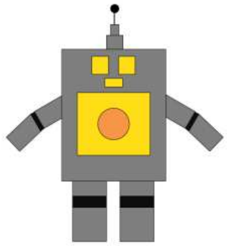
Press | Not Press | Either



Press | Not Press | Either



Press Not Press Either 

Press Not Press Either 

4. Do you have any other comments about the study? Do you have any suggestions to improve the instructions or the task itself to make it easier to learn?

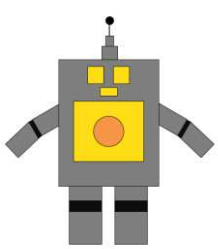
APPENDIX B

Reversal Paradigm Post-Task Survey

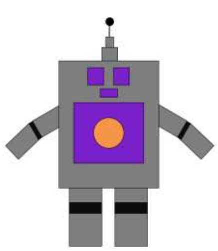
Please complete all questions in the order presented and avoid returning to previous questions.

1. Please describe the strategies you used, if any, to learn the correct response for ed to answer each robot

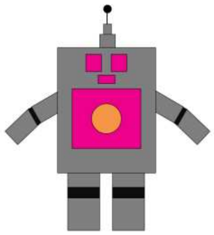
2. Before the robots suddenly changed they were followed by a win if you made the right response, or a loss if you made the wrong response. For this question, please disregard what the best response was (whether you think you should press the space bar, or refrain from pressing the space) and just let us know whether the robot was usually associated with winning or losing by circling the word "win" or "loss" If you think that the robot was associated with both winning and losing at times, you can circle the word "both". If you don't remember, circle "Don't know". We'd like to know which robot was associated with what, before the changes in the last round.



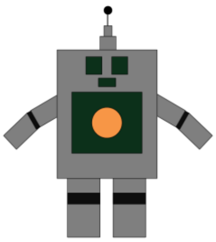
Win Loss Both Don't know



Win Loss Both Don't know

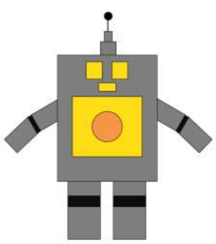


Win Loss Both Don't know

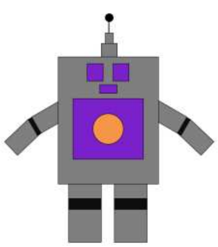


Win Loss Both Don't know

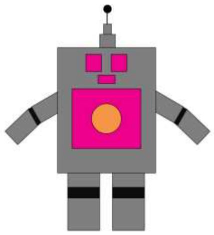
3. After the robots suddenly changed, they were followed by a win if you made the right response, or a loss if you made the wrong response. For this question, please disregard what the best response was (whether you think you should press the space bar, or refrain from pressing the space) and just let us know whether the robot was usually associated with winning or losing by circling the word "win" or "loss" If you think that the robot was associated with both winning and losing at times, you can circle the word "both". If you don't remember, circle "Don't know". We'd like to know which robot was associated with what, after the changes in the last round.



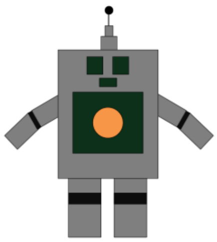
Win Loss Both Don't know



Win Loss Both Don't know



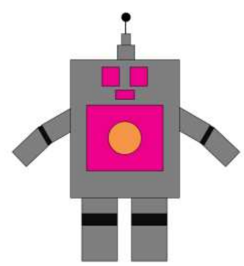
Win Loss Both Don't know



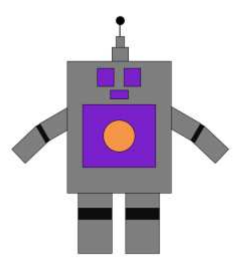
Win Loss Both Don't know

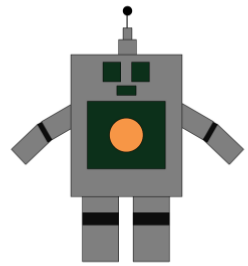
4. Before the robots suddenly changed, for some robots, the best response was to press the space bar (to win, or to avoid losing), whereas for other robots the best response was to not press the space bar. For this question, please disregard whether the robot was associated with a win or loss, and just indicate whether the best response is to press the space bar or not press for this robot. If you think the two responses are equally good, you can choose the word 'Either'. Circle the best option.

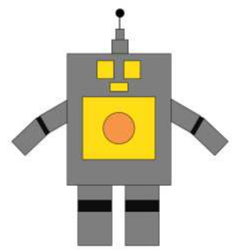
Press Not Press Either



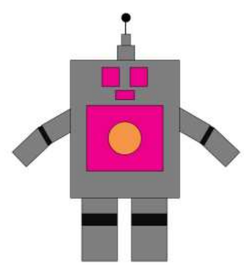
Press Not Press Either



Press Not Press Either 

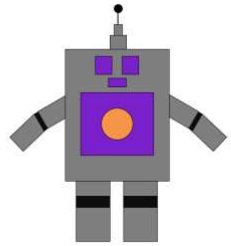
Press Not Press Either 

5. <u>After the robots suddenly changed in the last round</u>, for some robots, the best response was to press the space bar (to win, or to avoid losing), whereas for other robots the best response was to not press the space bar. For this question, please disregard whether the robot was associated with a win or loss, and just indicate whether the best response is to press the space bar or not press for this robot. If you think the two responses are equally good, you can choose the word 'Either'. Circle the best option.

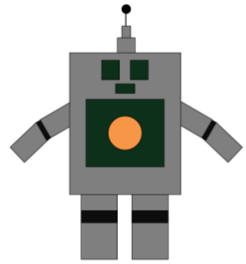


Press Not Press Either

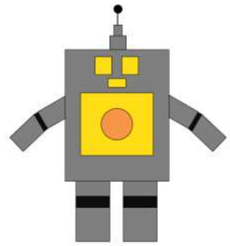
Press Not Press Either



Press Not Press Either



Press Not Press Either



6. Do you have any other comments about the study? Do you have any suggestions to improve the instructions or the task itself to make it easier to learn?