

Master thesis on Cognitive Systems and Interactive Media

Universitat Pompeu Fabra

Investigating Decision Making Process in a 2D Grid World: A Hotel Lobby Experiment

Cansu Ak

Supervisor: Berkay Demirel

Co-supervisor: Martí Sànchez-Fibla

September 2023



**Universitat
Pompeu Fabra**
Barcelona

Master thesis on Cognitive Systems and Interactive Media

Universitat Pompeu Fabra

Investigating Decision Making Process in a 2D Grid World: A Hotel Lobby Experiment

Cansu Ak

Supervisor: Berkay Demirel

Co-supervisor: Martí Sànchez-Fibla

September 2023



**Universitat
Pompeu Fabra**
Barcelona

Table of Contents

1.	Introduction	1
1.1.	Dual System In Human Decision Making	2
1.2.	Reinforcement Learning And Human Decision Making	4
1.3.	Effects of Video Game Playing on Decision Making.....	6
1.4.	Related Work on Habituation and Deliberation.....	7
1.5.	Navigating Habituation and Goal-Directed Arbitration: Anomalies in Contexts of Limited Observation	8
1.6.	Behavioral Paradigms of Habituation and Deliberation - Hotel Elevator Row Task	10
1.7.	HER - 2D Grid World Experiment.....	12
2.	Methods.....	15
2.1.	Participants.....	15
2.2.	Setup & Data Collection	15
2.3.	Experimental Set-up and Time Horizon	16
2.4.	Assessment of Habituation and Deliberation.....	18
2.5.	Data Analysis Method.....	19
3.	Results	21
3.1.	Trial Duration and Habituation.....	21
3.2.	Habituation failure and Re-habituation.....	23
3.3.	Trial Duration in Habituation, Common and Uncommon Trials.....	25
3.4.	Corridor choice	26
3.5.	Questionnaire	27
4.	Discussion	29
4.1.	Trial Duration and Habituation.....	29
4.2.	Habituation failure	29
4.3.	Corridor choice	30
4.4.	Limitations and Future Outlook.....	31
4.5.	Conclusion	34
5.	Bibliography.....	35
6.	Appendices.....	39
6.1.	Questionnaire	39
6.2.	Notebook for Data Analysis.....	40

Acknowledgments

I am expressing my sincere gratitude to my supervisor, Berkay Demirel, for his valuable advice and guidance, which were crucial in making this study a reality. I also wish to express my appreciation to Martí Sánchez-Fibla for his assistance and support throughout this thesis.

Abstract

Recent research into video game playing (VGP) has produced varied and inconclusive results. Past investigations have indicated potential positive impacts of VGP on cognitive abilities, specifically enhancements in areas such as visuospatial aptitude, processing speed, working memory, the ability to handle multiple tasks and speed of decision making. During the decision-making process humans operate with past experience and different cognitive skills when they take actions and make decisions in the world. In recent years there has been a wide range of research drawing on different fields such as psychology, neuroscience and the combination and interdisciplinary approach of these research fields in the context of artificial intelligence, specifically reinforcement learning. From early philosophy to the present, human behavior is anchored in the idea of operating with two minds. The field of neuropsychology have focused on two different modes of human decision-making process also referred to as the dual-process model of behavior: controlled versus automatic processes, one more deliberate interacting with goals and the other focusing on reflexive, habitual and driven by affects. Both modes habituation and deliberation have been subjects of research. The arbitration between both modes of the decision-making process has developed to a research interest. Due to biological and behavioral similarities in reinforcement learning, both processes can be examined and understood. The following thesis has the purpose of examining how video game experience influences habituation and deliberation processes in human decision-making.

Keywords: Dual-Process Model Behavior; Decision-making; Reinforcement Learning, Model-Free; Model-Based; Video-gaming; Cognition; Deliberation; Habituation

1. Introduction

In our everyday goal-directed tasks, humans have the tendency to focus on the relevant aspects of incoming sensory information from the external world and utilize them to transform it into motor actions. Human then harness this information to translate it into motor actions. At each phase of cognitive processing in the brain, these conversions involve selecting from various options to address questions related to what, where, when, and how. This intricate process constitutes a fundamental aspect of decision-making, requiring the combination of insights from diverse sub-processes such as sensation, perception, and action. These operations unfold within an extensive brain network encompassing multiple regions, as noted in studies by Gallivan et al. (2018) and Siegel et al. (2011, 2015).^{1 2} The swiftness and precision demanded by numerous perceptual and sensorimotor decisions remain a key factor.

Yet, the impact of enjoyable and mentally captivating undertakings like playing video games on the brain's composition and behavioral reactions within the realm of decision-making remains largely unexplored. The positive changes from these activities are important for training our thinking abilities. This training should help us become better at being correct or faster when we make decisions. Usually, getting faster means being less accurate, and being more accurate often means taking more time (speed-accuracy trade-off). But for these activities to significantly have an improvement for cognitive skills they should be able to improve humans to be both faster and more accurate at the same time, which is examined in the studies by Spieser et al. (2017).^{3 4}

As playing video games often requires quickly using information from senses and making accurate decisions over and over, it's a good way to study decision-making abilities. Studies have found that playing video games can improve working memory (Anguera et al., 2013), increased attention control (Anguera et al., 2013; Greenand and Bavelier, 2015) enhance the ability to see details in space and improved switching between different tasks (Oeiand and Patterson, 2014).^{3 5 6 7 8} Although the beneficial effects of video game playing and its effects on enhanced cognitive skills have been examined, performative underpinnings of these beneficial effects in decision-making regarding habituation and deliberation processes have yet to be explored.

In this study, we delve into the balance between these two systems, driven by insights from reinforcement learning and psychology. To achieve this, we employ a task known as Hotel Elevators Rows (HER, for brevity), which is designed within the Reinforcement Learning framework.⁹ The HER (Hotel Elevator Task) experiment involves creating a grid-based environment to explore and illustrate the cognitive dilemma that humans face when navigating through a hotel lobby, showcasing the interplay between two distinct decision-making systems. Our goal is to investigate the behavioral connections associated with habituation and deliberation. Given that a primary objective of this master's thesis is to conduct initial research into the suitability of the HER task as a behavioral tool for probing the fundamental aspects of the interplay between habitual and deliberate actions under the consideration of gaming experience.

1.1. Dual System In Human Decision Making

Current decision making process or dual process perspectives rely on two mechanisms of thinking processes, for example controlled versus automatic processes¹⁰, impulsive and reflective mechanisms¹¹ and System I and II¹⁰. part of decision-making has been the main focus for formal modelling systems.¹⁰

The Dual-Process Model of Behavior in psychology as described for instance by Metcalfe and Michels's hot/cool model⁸ differentiate between a "hot emotional system" and a "cool cognitive system" which determines a person's current state behavior and action. The MODE model of Fazio and Towles-Schwen's has a similar approach of distinguishing the two systems: spontaneous and deliberative, automatic behavior guiding spontaneous actions and explicit behavior guiding deliberative actions. According to Kahneman, these two mechanisms are classified as System 1 and System 2.¹⁰

The first system is mainly associated with a intuitive, automatic and rapid decision-making process with minimal effort. The habitual system or behavior in human decision-making is gained through experience and accumulated knowledge of the environment, which holds its base in trial and error. Briefly, this system is intuitive rooted in expected events.^{12 13}

The second system, a goal-directed system in the decision-making system is rather flexible and selects within different options, which indicate that this behavior is sensitive towards the environment.¹⁴ System 2 demands higher cognitive power and attention.¹⁰

However, these mechanisms come at a cost in terms of effort and energy expenditure. Specifically, within the context of the analyzed arbitration process, attention assumes a crucial role. It serves as a bottleneck, managing the overwhelming influx of information that constantly bombards the brain. Moreover, attention is triggered when habitual behaviors fail to achieve their intended goal, resulting in habituation failures. In such cases, more resource-intensive mechanisms come into play to analyze what went wrong and resolve the issue.⁹

In these scenarios, the individual's cognitive model of the world, shaped by the mind's innate ability to identify patterns and rules in the environment, may not align with the actual world. Habituation failures can stem from various factors. For example, they might occur when attempting to derive rules from a chaotic environment, rendering the task impossible. Changes in the environment, due to their rapid and inherent nature, can challenge the habitual system's ability to adapt quickly.¹²

Furthermore, the inherent complexity of our world can lead to habituation failures. Our awareness of the environment is inherently partial, and its unpredictable randomness adds another layer of complexity. Another contributing factor is the brain's inclination towards efficiency, often opting for the least effortful and quickest option, even if it may not be the most suitable response for the situation. This is closely linked to the concept of the speed-accuracy trade-off, well-documented in fields like reinforcement learning and cognitive neuroscience. This trade-off involves a tendency to prioritize decision speed, potentially compromising decision accuracy.¹⁵

Within this thesis, the focus lies on habituation failures arising from the intricacies and randomness of our environment, as well as the brain's inclination towards efficiency. Central questions include how the brain navigates such scenarios and transitions from habitual to deliberate behaviors to address anomalies. Additionally, the thesis aims to uncover the behavioral manifestations of this arbitration process, while examining if VGP have an overall enhanced performance than Non-Video Game Players (NVGP).

1.2. Reinforcement Learning And Human Decision Making

It is essential that we adjust our worldview after errors occurred, in order to react to the surroundings, the most convenient way. For learning, behavior modification, and maximizing future rewards in any scenario with significant consequences, it is essential to comprehend how humans utilize information about mistakes (henceforth "feedback") to later adjust our behaviors. However, it is not well known how humans react when error happens in particular, whether any changes in behavior that follow are necessary conscious or explicit, especially within the field of gaming and its' cognitive performances.¹⁶

Since reinforcement learning has the most similar learning structure to that of humans and other animals, it may be used to explain a variety of issues in psychology and neuroscience. Particularly many fundamental reinforcement learning algorithms take their cues from biological learning systems. For instance, it has been proposed that Temporal Differences Learning can replicate the "reward prediction error hypothesis of dopamine neuron activity" in reinforcement learning.¹⁷

Two primary forms of reinforcement learning can be distinguished - Model-free and Model-based - and this division can be likened to the System 1 and System 2 distinction discussed earlier. Model-free algorithms (like Q-learning), on one hand, learn through experience by developing a policy. They operate without constructing a model of the environment, resulting in quick adaptation but limited responsiveness to changes, as they link states and actions through the value function. This form of reinforcement learning is rooted in the conventional stimulus-response-outcome approach.¹⁸

Conversely, Model-based systems (like Dyna-Q) possess a model of the environment, allowing them to predict potential outcomes based on sequences of actions. Such systems make decisions by projecting actions forward, leading to slower immediate responses and learning. Constructing a model necessitates time; however, these systems can rapidly adapt to alterations in the environment. Although useful in research, rigid dichotomies like these oversimplify the nuanced nature of decision-making and learning. Researchers are now integrating various learning and decision-making models to gain a more comprehensive understanding.¹⁹

One emerging approach gaining prominence is the Successor Representation, initially formulated by Peter Dayan. While it was first conceived to enhance the generalization of Temporal Difference learning, it is now recognized as a valuable model for learning and decision-making, complementing both model-based and model-free approaches. Successor representation is a sophisticated temporal depiction of states, a forward-looking map detailing the anticipated future presence of successor states. In this framework, learning revolves around the expected frequency of visiting other states from a given starting point.^{19 20}

Computationally, Successor Representation mirrors the value function, but instead of tallying rewards, it aggregates state occurrences. Notably, it also shares an equivalence with temporal difference learning. While temporal difference learning employs reward prediction errors to update value estimates for state-action pairs, Successor Representation adapts this algorithm to measure state disparity errors - the difference between actual and predicted state occurrences. Moreover, Successor Representation brings a computational advantage by simplifying value computation, transforming it into a linear amalgamation of expected future state occupancies for each successor state, weighted by their expected rewards.²⁰

These computational features allow Successor Representations to facilitate multi-step decision-making, rendering them computationally efficient yet more prone to errors compared to model-based algorithms when confronted with changes in the environment's transition dynamics. This susceptibility arises because the Successor Representation encodes the environment's transition function into its predictive map, effectively erasing individual transitions. Consequently, it becomes blind to local transition modifications, necessitating relearning similar to how model-free algorithms adjust their value function during environmental shifts.²¹

The concept of Successor Representation models has also been proposed as a predictive map for spatial navigation, aligning with the cognitive map metaphor introduced by Tolman.²³ Tolman's research points out that the hippocampus plays a role in representing space by using place cells that link to various elements within the Successor Representation matrix. Additionally, thresholded eigenvectors of the Successor matrix are

linked to the activity of grid cells. Moreover, this perspective suggests that dopamine's function extends beyond reward prediction; it serves as a vector-valued error signal, utilized to revise and modify this Successor Representation.²⁴

1.3. Effects of Video Game Playing on Decision Making

In a study from Jordan and Dhamala (2022) behavioral responses were observed in which VGP exhibited quicker decision-making, with an approximate time advantage of 190 milliseconds, and greater accuracy by around 2% compared to NVGP. Turning the attention to brain responses, they investigated changes in brain activity within commonly activated regions between the two groups. This analysis revealed that video gamers demonstrated elevated task-related signal changes in specific brain areas. Notably, these areas included the right lingual gyrus, right supplementary motor area (SMA), left thalamus and hippocampus, creating a mental map of space, all of which correlated with the observed improvements in behavioral performance regarding the specific study. According to the study's findings, VGP individuals outperform NVGP counterparts not only in terms of response speed but also in terms of accuracy, challenging the notion that gaming experience hinders cognitive performance. The absence of a speed-accuracy trade-off suggests that engaging in video game playing could be a promising avenue for cognitive training, particularly in relation to enhancing decision-making abilities.²⁵

Early investigations into the potential positive impacts of video games on cognitive abilities included a series of experiments that strongly supported a positive connection between video game training and enhancements in spatial memory. Video game players (VGP) experienced improved visuospatial capabilities by enhancing their ability to memorize object locations, track objects, and perform mental rotations. In these early studies, gamers consistently outperformed non-gamers on various cognitive tasks. To establish a clearer cause-and-effect relationship, researchers went further by training non-gamers to play video games over a span of weeks. This training led to longitudinal improvements in the cognitive performance of these newly trained gamers, particularly in tasks involving task switching and object placement memorization.²⁶

These findings support the possibility of a causal relationship and counter the notion that individuals possessing these skills are simply drawn to playing video games. This, in turn,

paves the way for more in-depth investigations into how cognitive skills might transform following engagement with video games.

Beyond the initial research by Green & Bavelier (2007), several subsequent studies have also identified robust connections between video game playing and advancements in visuospatial skills, encompassing areas like visual attention, object tracking, visual memory, and task switching.^{26 27}

Beyond the enhancements observed in visuospatial skills, researchers delved into the correlation between video game engagement and more intricate cognitive abilities, such as problem solving. As stated by Hamlen (2012), adept game players have demonstrated elevated aptitude in activities like seeking information, categorization, risk assessment, strategizing, critical thinking, and having confidence in their knowledge. The author also outlined a range of potential skills and strategies employed in game playing that encompass these types of problem-solving behaviors. This implies that the effectiveness of learning from game playing could potentially extend to other contexts.²⁸

As researchers investigated deeper into the impact of video games on cognitive abilities, subsequent investigations consistently revealed robust connections between video game engagement and executive control skills. These skills encompass multitasking, the ability to divide attention, task-switching, processing speed, working memory, and enhanced reaction time without compromising accuracy. An illustrative instance of this is the skill of split attention or multitasking, which proves valuable when studying amidst distractions like television, phone calls or messages or ongoing conversations beneath (cognitive load).⁴

Moreover, within the context of gaming, rapid responses are often essential, rewarded, and even encouraged. This emphasis on swift reaction times can potentially yield advantages in swiftly and accurately responding to tasks, thereby fostering fluency. Consequently, these improvements appear likely to positively influence performance in timed tests, creating a favorable impact on timed test results.²⁹

1.4. Related Work on Habituation and Deliberation

Although gamers cognitive skills enable them to react fast and more accurate and have an overall better performance in solving tasks and making decision, often as mentioned

before we as humans are under partial observability and this leads us to make decision either based on habituation or deliberation (System 1, System 2).

Consequently, numerous studies have employed tasks that indirectly differentiate between behavioral types by focusing on decisions that either indicate a discernible understanding of state relations or lack such understanding under specific conditions. This discernment typically resides within model-based systems, which, as previously highlighted, align with deliberate behavior.

A substantial body of research into decision-making, particularly involving non-human animals like rodents, has been conducted. Rodent studies, often involving foraging tasks, provide valuable insights as animals navigate decisions regarding resource allocation, such as time management. Through research, observations shows a shift from deliberate to habitual behavior as individuals gain experience, as well as a shift from habitual to deliberate behavior in situations with increased decision complexity.^{30 31}

Within the realm of rodent studies, the concept of Vicarious Trial and Error (VTE) has gained prominence. VTE is a spatial behavior exhibited by rodents, characterized by forward and backward contemplation of possible trajectories at decision points, frequently accompanied by pauses. This behavior is often associated with exploration within an internal mental space of potential outcomes—a cognitive map of how the environment operates.³² It emerges once a cognitive map of the task is established and dissipates once action sequences become automated, signifying the onset of habituation. VTE is seen as a manifestation of indecision during deliberation in rodents and has even been linked to human behavior and decision-making.^{33 34}

1.5. Navigating Habituation and Goal-Directed Arbitration: Anomalies in Contexts of Limited Observation

In the experiment HER (Hotel Elevator Task) a grid environment is designed to explain the human behavioral challenge between the dual systems as a navigational task. HER is a grid illustration of a hotel hall, which explores habituation failures that occur, due to changes in the starting point of the environment. The work elaborates on how human

investigate and cooperate with an anomaly in the environment and rehabilitate again in the now transformed environment.⁹

In this research line, the aim is to further elaborate and gain more clarity in the arbitration between two main human decision-making of habitual and goal-directed behavior by creating habituation-based environments, habituation failures due to environmental transformations.

In the study by Neto (2022), an extension is made upon the previous investigation conducted by Sanchez-Fibla (2020). This expansion involves a transition from a grid world reinforcement learning environment experiment to a human behavioral experiment utilizing virtual reality (VR) technology experiment.

Furthermore, the research conducted by Neto (2022) yields human behavioral data that is comparable to the work previously mentioned. Consequently, this section will delve into comprehensive details regarding the methodology, outcomes, and ensuing discussions related to Neto's research.

The objective of the human behavioral experiment centers on elucidating the underlying dynamics involved in the arbitration between habitual and goal-directed behaviors within the context of the HER task, as initially introduced by Sanchez-Fibla (2020). To achieve this, a Virtual Reality (VR) experiment involving fifteen human participants was undertaken and achieved to show habituation failures due to environmental changes and mechanism of human decision making that has been mentioned in the literature above.

The original study of Sanchez-Fibla has also been further investigated in a study of Gezer (2022). Within the scope of the study, the primary objective was to enhance the understanding of the intricate interplay between two core facets of human decision-making: habitual and goal-directed behaviors. To achieve this, Gezer adopt a reinforcement learning framework that encompasses modeling the process of habituation, failures in habituation due to environmental alterations, and the subsequent rehabilitation process for these changes. Additionally, the study intended to explore the efficacy of the proposed alarm mechanism featured in the Successor Representation-monitoring model-free behavior introduced by Sanchez-Fibla, particularly in a 3D setting.

To accomplish these objectives, a 3D rendition of the Hotel Elevator Row (HER) environment, an extension of the original grid-based version introduced in Sanchez-Fibla

was transformed. The core aim underlying the design of the 2D HER environment was to simulate habituation failures through shifts in the starting point of a solitary learning agent. This agent is subjected to training employing two distinct approaches: model-free reinforcement learning using deep Q-learning (DQN), and Successor Representation (SR). These techniques were employed to address a room-finding task within the 3D Hotel Elevator Row (HER) environment.

1.6. Behavioral Paradigms of Habituation and Deliberation - Hotel Elevator Row Task

In essence, the HER task serves as a means to identify deviations in habituation patterns among human participants. The task unfolds within an environment resembling a hotel setting, where participants embark from the hotel lobby corridor situated on the ground floor. Notably, two elevator doors stand opposite each other – one on the left and another on the right – ferrying participants to the room floor. Each trial commences with the opening of a single elevator door, be it the left or the right. Such trials are categorized as common or uncommon, depending on whether the left or right elevator door is opened.

Further within the environment, the room floor features a dozen symmetrically arranged hotel room doors. These rooms are positioned in a four-sided, symmetrical layout. This HER task essentially constitutes a one-choice, partially observable task, embedded within the framework of Reinforcement Learning, specifically a Partial Observable Markov Decision Process (see Figure 1).

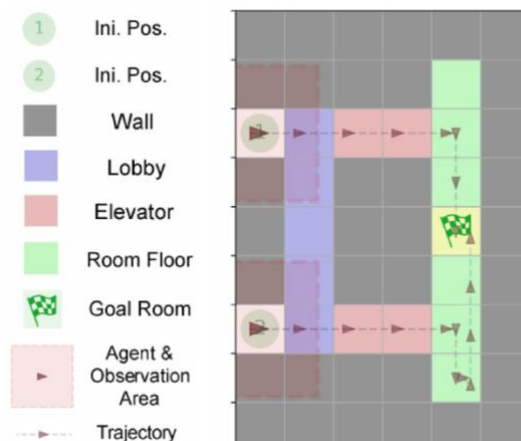


Figure 1: HER Task Environment by Sanchez-Fibla (2020)

This unique hotel environment deviates from convention by having elevators positioned opposite each other in the lobby, as opposed to the typical side-by-side arrangement. A play of circumstance may lead individuals to consistently opt for the common elevator on the right side of the corridor while consistently overlooking the left (uncommon) elevator. However, a chance encounter with the uncommon elevator – if it happens to be open – might catch one off guard, prompting the automatic sequence of actions typically associated with reaching one's room. For instance, entering the elevator, turning left on the designated floor, and once more turning left to locate the room. Unbeknownst to the individual, they have made a habituation failure. The fundamental premise of the HER task revolves around the notion that investigating such failures in habituation can offer valuable insights into the dynamics between deliberation and habituation.

In the context of reinforcement learning, the agent's goal is to efficiently navigate from the hotel lobby to a specific room as quickly as possible. It's important to note that there are two potential starting points, each with varying probabilities of occurrence: a common one (Initial Position 1) and an uncommon starting point (Initial Position 2).

In the VR experiment of Neto (2022), participants navigate in various directions and interact using a controller. They aim to reach room 303 in each trial. The experiment consists of three phases: an initial exploration phase, a habituation phase (trials 2-12), and a testing phase. Uncommon trials (13, 18, 21, and 28) (see Figure 3) in the testing phase introduce anomalies by opening the left elevator door on the first floor, challenging participants' habits.¹⁴ (see Figure 2)

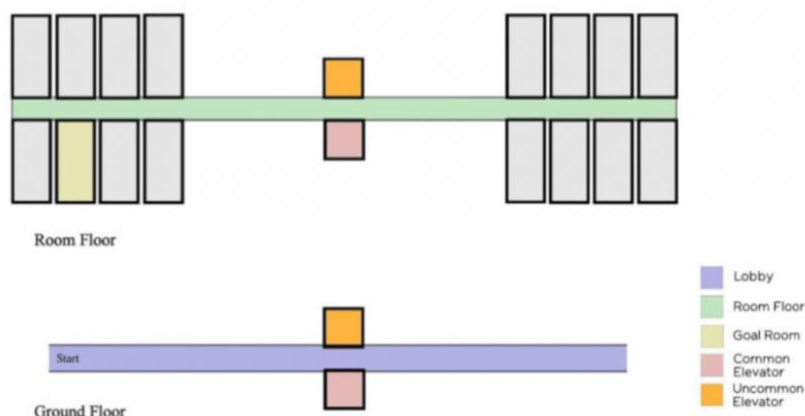


Figure 2: HER VR Environment by Neto (2022)



Figure 3: VR Experimental Setup

With the successful demonstration of habituation failures in Neto's (2022) VR experiment, which offers enhanced ecological validity, this thesis shifts focus to a 2D grid world setup. This new setup is designed to assess gamers' motor and cognitive abilities using keyboard inputs, aiming to validate the conjectured associations between gaming experience and overall performance, showcasing improvements over non-gamers.¹⁴

1.7. HER - 2D Grid World Experiment

With the successful demonstration of habituation failures in Neto's (2022) VR experiment, which offers enhanced ecological validity, this thesis shifts focus to a 2D grid world setup. This new setup is designed to assess gamers' motor and cognitive abilities using keyboard inputs, aiming to validate the conjectured associations between gaming experience and overall performance, showcasing improvements over non-gamers.

Participants in the HER behavioral experiment will be asked to navigate as quickly as they can to their room in a 2D environmental setting (Figure 3). The participant first enters the hotel, then navigates from the lobby to the elevators, and finally exits the elevator on the correct floor and then finds the correct path to the asked room. The first time the trajectory or navigation is still deliberate (model-based, system 2), requiring the most attention, as the environment is new. After experiencing the hotel task multiple times, participants become habituated, enabling them to reach their designated room with minimal environmental attention and freeing up cognitive resources for other tasks.

In addition, since in this experiment, the hotel has the special case of having elevators on opposite sides of the lobby, instead of side by side, it creates a proneness of habituation failures. The thesis suggests that when one arrives to a certain level of the habituated behavior, they will by default turn to the side they became habituated to. The common elevator will open more, causing habituation, while the rare elevator will open on occasion.



Figure 4: HER Task - 2D Grid World

The experimental task has different decision points similar to Neto (2022)¹⁴: (1) stepping into the elevator, which is the crucial moment at which the participants can gather the information needed to account for changes in the environment at rare trials; (2) exiting the elevator, which is the main decision point at which the subjects will need to choose which is the best corridor to reach their room – by turning in either two direction, left or right; and (3) there is the final decision point, at which the participants must decide whether to turn left or right. Three sections will make up the behavioral experiment. The experiment consists of a habituation and testing phase. In the habituation phase participants are instructed to arrive in the specified room as quickly as they can. In the habituation phase, the participant gets habituated to the environment, in order to create a mundane effect of the experiment. The common elevator will open most of the time, while the uncommon elevator will open at predefined intervals. These relevant trials are used to investigate habituation failures in the experiment. The second component is the testing phase. Since first most of the time (common trials) the participant will take the same path

to the room (turning right after exiting the elevator in a symmetrical environment regardless of which elevator has been exited, and secondly the only cue to indicate which elevator is open is green or red. Put differently, the thesis suggests that during their initial experience in the hotel, individuals prioritize the primary task of finding the main path, but as they become habituated through successive trials, their attention to this task gradually diminishes.

The thesis hypothesizes, that VGP are faster in the habituation phase than NVGP, due to research claimed cognitive skills.

Another hypothesis is that Gamer's performance to find the correct room in uncommon trials (deliberation) is greater than with Non-Gamers, and Gamer's performance to find the correct room in common trials (re-habituation) is greater than with Non-Gamers, hypothesizing that VGP show overall better performance in learning from habituation failures and implementing solutions faster to errors than NVGP.

2. Methods

2.1. Participants

The study encompassed twenty-six participants, including 13 females and 13 males, with an average age of 24 years falling within the range of 21 to 33 years. Participants were students at Pompeu Fabra University (UPF) in Barcelona, students at the University of Vienna and employees of a video game company in Vienna. All of them volunteered for the experiment without receiving any kind of compensation. Initially, the participant count was twenty-six, but two were outliers due to high time durations.

Participant selection followed stratified convenient sampling, adhering to specific criteria: proficiency in English, gaming experience, within the age range of 18 to 50, possessing normal or corrected-to-normal vision, lacking arm-related trauma or hand difficulties.

Prior to participation, participants were informed about the study's general purpose and provided their consent. Despite the relatively small sample size comprised of university students, the study's first approach nature and its exploration of a topic common to humans over time minimize the potential impact of this limitation on result reliability.

2.2. Setup & Data Collection

The experiment unfolded predominantly at the Poblenou Campus of UPF and occasionally at the University of Vienna Campus. Participants were engaged in the study using laptops, upon which the experimental environment was meticulously established. Navigation through this environment was facilitated using three specific keyboard keys. A questionnaire was handed out to assess each participants gaming experience with a Likert scale form, from “very experienced” to “very inexperienced”. Notably, the original virtual reality setup from Neto's (2022) work was crafted within Unity and subsequently adapted into a 2D Grid World framework. The resulting data, encompassing various variables outlined were diligently logged and subsequently organized into a comprehensive CSV file for further analysis.

2.3. Experimental Set-up and Time Horizon

Data collection was conducted within a single session for each participant. While this compressed timeframe may seem limiting, it's important to note that the subject matter under investigation involves a continuous interplay between habitual and deliberate decision-making, a phenomenon recurrently experienced in our daily lives. Moreover, the rapid occurrence of habituation lends credence to the notion that the compressed data collection sessions are unlikely to exert a notably adverse impact on the robustness of our results.

The experiment unfolds as a single session lasting approximately 10 minutes per participant, encompassing a path-finding task termed the HER task, coupled with a concise concluding questionnaire (refer to Appendix 2). More specifically, participants find themselves within a hotel lobby, tasked with expeditiously reaching their designated hotel room while attentively managing notifications on a virtual mobile phone.

The experiment comprises two distinct segments, depicted in Figure 5. Firstly, a habituation phase, which allows participants to acclimate to the environment and navigation mechanics. During this phase, the primary objective is swift arrival at their hotel room. Notably, the elevator that consistently opens during this phase is the one positioned on the right side of the corridor, labeled as the common elevator. The habituation phase, spanning 6 trials, maintains the environment's consistency. Once the correct room is found, the trial ends automatically and the participant finds him/herself again in the hotel lobby.

In the final phase, the testing phase, encompassing 16 trials, an unusual anomaly is introduced by having the left-side elevator (the uncommon elevator) open in every fifth trial. This anomaly triggers a HER anomaly, representative of a habituation failure. This phase comprises four sets of five trials each, commencing with an uncommon trial followed by three rehabilitation trials involving the common elevator. The last set comprises a single uncommon trial, as rehabilitation is not required.

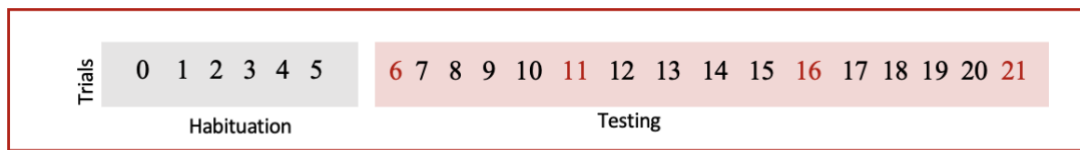


Figure 5: Trials of experiment. Uncommon trial numbers are marked dark red.

During the environment design phase, we encountered a series of choices aimed at achieving the HER anomaly while preserving spatial information regarding room positions. The original paper's approach in Sanchez-Fibla (2020), where artificial agents learned solely through trial and error without spatial clues, was suited for a human behavioral experiment, especially due to the design decision of opting for a 2D grid environment. As the environment was limited in the design, no cues as to where the room is located were given.

The environment was deliberately designed to maintain symmetry, regardless of whether participants exited from the left or right elevator. This design choice aimed to prevent the task from becoming overly simplistic for participants, ensuring that the room-finding challenge remained consistent. To achieve this, a uniform color scheme was implemented throughout the environment, featuring white doors and a grey corridor. Participants navigated through the environment using keyboard controls (up, left, and right). This decision was made to uphold the symmetrical nature of the environment, as the study did not leverage the advantages of a 3D virtual reality setting as observed in Neto (2022), nor did it utilize a solely agent-based environment like in Gezer (2022). Despite these differences, the central objective was to examine whether the spatial and speed accuracy trade-off observed in gamers was superior to that of non-gamers. By maintaining symmetry and uniformity in the environment, the experiment sought to isolate the impact of gaming experience on performance. This design ensured that any differences between gamers and non-gamers could be attributed to their respective experiences rather than the layout of the environment itself.

2.4. Assessment of Habituation and Deliberation

The trial duration at the beginning of the experiment will likely be longer as participants familiarize themselves with the environment, requiring time to navigate, trigger wall signs for room location information, and identify rooms. As participants gain experience and develop a cognitive map of the environment, this initial need for exploration diminishes, resulting in faster trial durations. Consequently, trial duration indirectly reflects participant speed. However, uncommon trials demand more deliberation due to the adjustment needed to reach the desired room without detours. This deliberation, associated with indecision and hesitation, leads to longer trial durations. The rehabilitation trials subsequently counterbalance this effect. Hence, trial duration serves as a comprehensive measure to discern the dominant decision-making system: lower values signify habitual behavior, while higher values suggest deliberate behavior.

To compute trial duration, we initially converted timestamps into seconds by subtracting each timestamp from the initial timestamp and take the difference for further analysis. Then, the difference between the trial end time and trial start time was calculated. For habituation, a Mann-Whitney-U test was employed across all trials to explore any correlation between trial duration and the increasing number of trials. For deliberation, also a Mann-Whitney-U test was used to determine if common and uncommon trials.

Our hypothesis centered on the potential extraction of decision-making systems through secondary data indicators, including participants' frequency of consulting signs for clues in each trial and its evolution over time, as well as the frequency of errors, such as entering the incorrect corridor. By analyzing these aspects, we aimed to discern the prevailing decision-making system. Our examination encompassed trajectory data as steps transformed into action data to analyze which corridor has been taken in all the trials, but especially in the uncommon trials to look for habituation failures.

Regarding error analysis, we initially categorized errors based on the chosen corridor, meaning which step has been taken when the elevator has been exited. Subsequently, we assessed the total number of errors for the false corridor type depending on the exiting elevator. Additionally, we computed the count of mistakes made in each trial per false corridor error. To ensure accuracy, we considered only the first instance of step taken to

the right direction for each participant within a trial, excluding subsequent steps after that from the analysis was crucial to prevent skewed results.

2.5. Data Analysis Method

Utilizing Python and its libraries for data analysis, we examined behavioral patterns derived from collected data, with a focus on changes in reaction times, time duration steps and actions.

The principal objective of this master thesis is to initiate exploratory research into the viability of the HER task as a behavioral tool for investigating the fundamental intricacies of the interplay between habitual and deliberate behavior. Given the limited research on behavioral correlates associated with this arbitration, our goal is to ascertain whether the HER task is a suitable platform for advancing our comprehension in this domain, more specifically if it can provide into trained cognitive skills in our example gaming experience to help us investigate this field. In Neto (2022) this involved addressing two key inquiries: (1) can the study induce habituation in human participants using the VR task, and (2) can the study induce a habituation failure, or HER anomaly, in the HER task, along with identifying the corresponding behavioral patterns. As Neto (2022) showed in her thesis the HER anomaly did occur, within this thesis the important question is if the HER anomaly can be induced in people with gaming experience and test their spatial cognitive skills as well as their overall performance.¹⁴

To begin, we sought to identify the behavioral markers of habituation and validate its occurrence within the context of the HER task. Given the task's relatively straightforward nature and compact design, we anticipated that habituation would manifest swiftly, primarily due to humans' adeptness in navigating paths. A pilot data analysis further supported this notion, with habituation indicators such as trial duration stabilizing around 19 seconds per trial after approximately 6 habituation trials. Due to the simplistic and mundane environment we only opted for 6 habituation trials and not 8 as initially planned, as the task is repetitive.

Moving on, we contended that the behavioral patterns indicative of habituation would be counteracted by the emergence of the HER anomaly. Participants are more compelled to

focus on the steps needed to reach the designated room, potentially resulting in hesitation or indecision while deliberating and seeking external cues, nonetheless we hypothesize that the HER anomaly triggers the participant's realization that something is amiss, prompting efforts to pinpoint the issue and devise a solution, when the uncommon trials appear. Consequently, our hypothesis centers on the exploration of correlations encompassing both habituation failures and deliberative behavior within the field of gaming experience.

Following the 2D grid world experiment's conclusion, we administered a brief questionnaire (refer to appendix 3) to gather both general demographic information and experiment-specific data. In addition to inquiring about participants' gaming experience as this was crucial to determine gamers and non-gamers for further analysis, using a likert scale measure, gender and age, we sought to glean insights into other factors that could be pertinent to the experiment, which were the following:

Experience of errors: To gain insight into participants' experiences and to determine whether they detected the HER anomaly, we posed a non-influential general question. Participants were asked if they encountered any instances of mistakes or errors during the experiment, and if they did, whether they employed any strategies to address them.

Solution strategy: To gain insight into participants thinking pattern and strategy they used to solve any kind of mistake they encountered, as this serves as a better understanding of possible cognitive maps or other cognitive skills.

3. Results

3.1. Trial Duration and Habituation

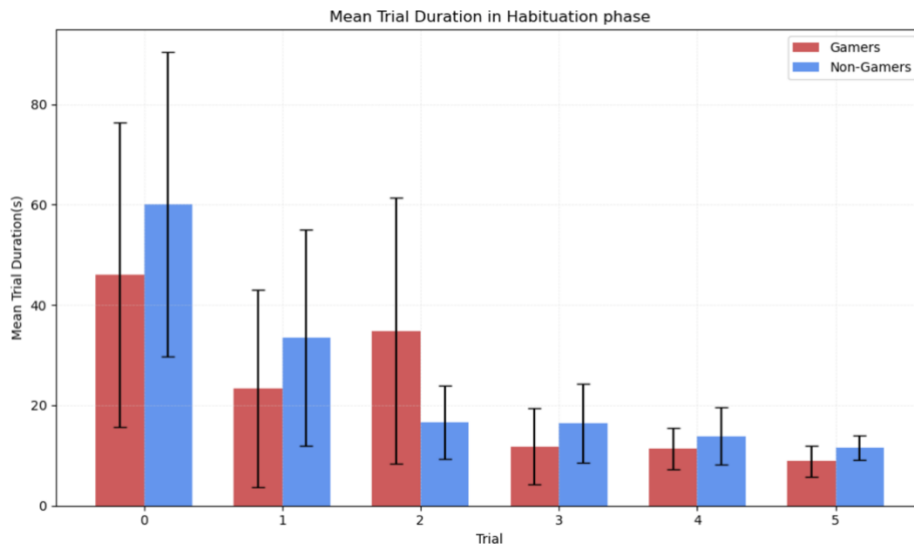


Figure 6: Mean trial duration for habituation phase with standard deviation. Gamers mean duration=13,96 s with standard deviation=26,58 s; Non-Gamers mean duration=19.63 s with standard deviation=30.01 s; mann-whitney-u test $p=0.001$, therefore $p<0.05$

The provided graph displays the mean trial durations along with their standard deviations for the habituation phase of the experiment, with separate bars for Gamers and Non-Gamers. Each bar represents a trial within the habituation phase, numbered from 0 to 5. The x-axis corresponds to the trial number, while the y-axis represents the mean trial duration in seconds. From the graph, we can observe the following:

Gamers vs. Non-Gamers Mean Duration: Gamers had a lower mean trial duration (13.96 seconds) compared to Non-Gamers (19.63 seconds) throughout the habituation phase. This difference suggests that Gamers generally completed each trial faster than Non-Gamers. The standard deviations of both Gamers and Non-Gamers show a range of variability in trial durations, as indicated by the vertical error bars representing the standard deviations. The longer error bars for Non-Gamers imply greater variability in their trial durations compared to Gamers.

The results of the Mann-Whitney U test ($p=0.001$) suggest a statistically significant difference in trial durations between the two groups, the p-value being less than 0.05.

In trial 2 we can see an anomaly, an additional observation provided is that in trial 2, Gamers took slightly longer than Non-Gamers. This variation in trial duration in trial 2 can be attributed to a specific participant among the Gamers group who spent more time orientating themselves within the virtual environment. This elongated orientation phase led to an outlier that impacted the mean duration for Gamers in trial 2.

In summary, the graph illustrates the differences in mean trial durations between Gamers and Non-Gamers during the habituation phase. Gamers, in general, completed trials faster than Non-Gamers, leading to a statistically significant difference. The longer duration in trial 2 for Gamers is explained by the presence of an outlier participant who took more time to orient themselves within the virtual environment.

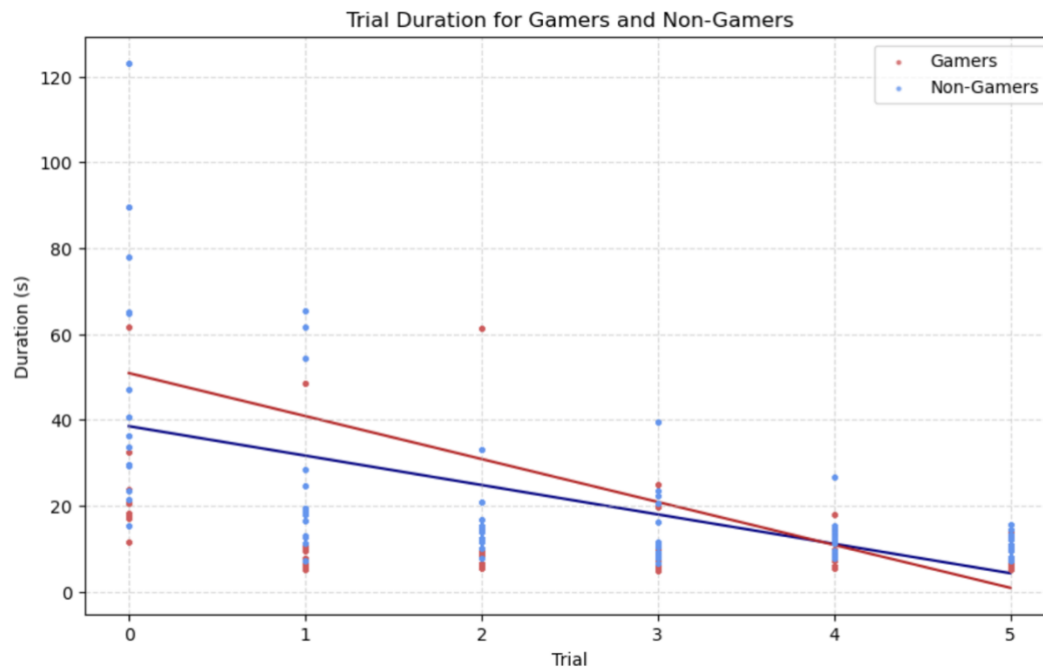


Figure 7: Linear regression model for trial duration between Gamers and Non-Gamers. Gamers: $r^2 = 0.245$; Non-Gamers: $r^2 = 0.387$;

The graph seems to depict a comparison between Gamers and Non-Gamers in terms of trial duration over the first six trials of the habituation phase. Looking at the data points and regression lines, you can infer the following: For both Gamers and Non-Gamers, the trial duration seems to decrease as the trial number increases. This suggests a potential habituation effect where participants become more familiar with the task and perform it more efficiently over time. The overall trend shows that within each group, it appears that the trial duration decreases at a similar rate for both Gamers and Non-Gamers, as

indicated by the negative slope of the lines. The significant difference in the starting points of the two groups suggests that there might be a baseline difference in performance between Gamers and Non-Gamers before the habituation phase. The graph shows the expected results of the hypothesis that Gamers are faster in the first 6 trials in the habituation phase than Non-Gamers, but we can not say that they are faster in the habituation process itself, due to the fact that Non Gamers were significantly slower in the beginning.

3.2. Habituation failure and Re-habituation

In the following section we examined if Gamer’s performance to find the correct room in uncommon trials (deliberation/habitaution failure) and rehabituation is greater than with Non-Gamers. The graph plotted below displays the mean trial durations for Gamers and Non-Gamers in uncommon trials (6,11,16,21) where participants need to find the correct room (habituation failure test).

The mean and median values show that Gamers, on average, have lower trial durations than Non-Gamers, indicating faster performance. The difference between the means supports your hypothesis and aligns with the Mann-Whitney U test results.

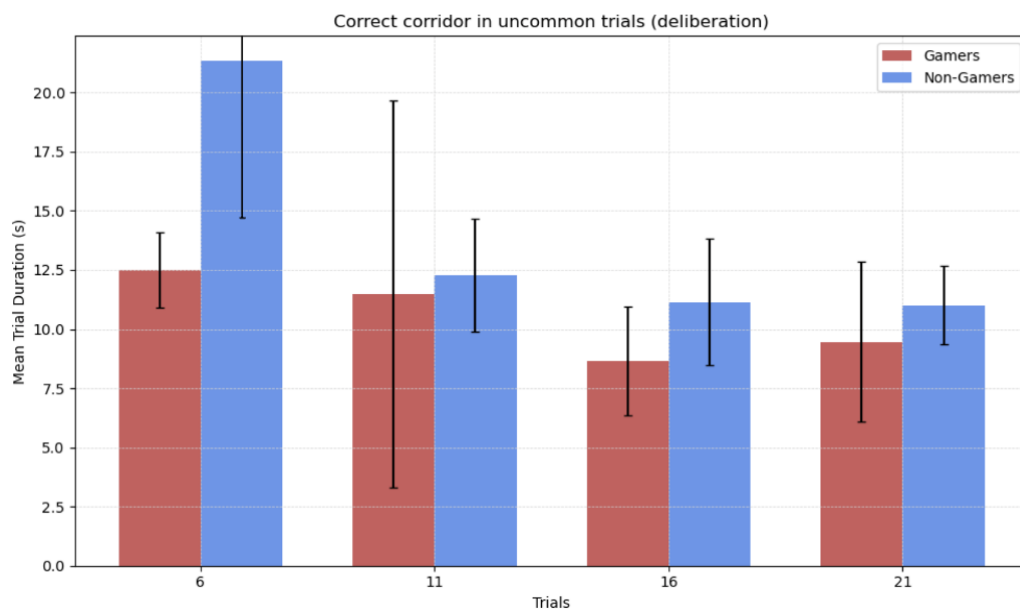


Figure 8: Mean trial duration for Gamers and Non-Gamers in uncommon trials; $p = 0.00715$

Additionally, the lower standard deviation for Gamers (2.89) compared to Non-Gamers (3.45) suggests that the trial durations among Gamers are less variable, indicating more consistent performance in finding the correct room.

Most of the participants as expected encounter their first habituation failure at trial 6, when the

switch of the elevator happens for the first time. After that both Gamers and Non-Gamers failure decrease. In trial 16 Non-Gamer already have a significantly lower failure than in trials 6 and 11.

We also hypothesized that Gamer’s performance to find the correct room in common trials (re-habituation) is greater than with Non-Gamers.

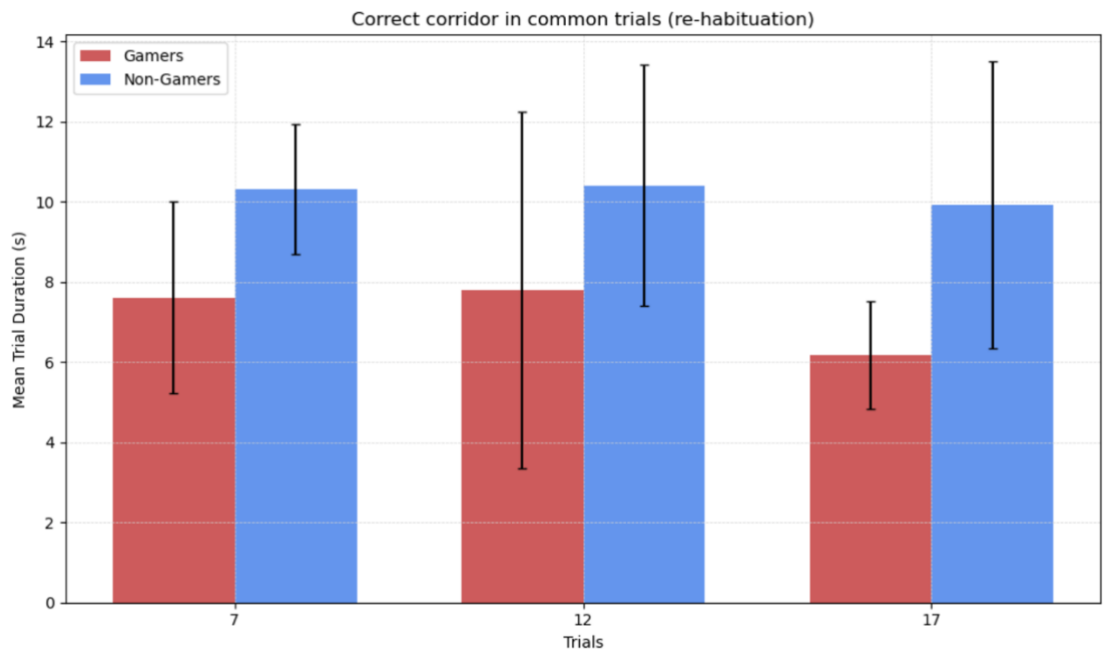


Figure 9: Mean trial duration for Gamers and Non-Gamers in common trials (rehabilitation); $p = 0.001$

The graph/p-value from the Mann-Whitney-U test confirms the hypothesis that Gamers take the correct corridor in common trials faster. However, Non-Gamers have a small difference on trial duration, about four seconds, as the median comparison shows.

Additionally, the lower standard deviation for Gamers (2.51) compared to Non-Gamers (3.12) suggests that the trial durations among Gamers are less variable, indicating more

consistent performance in finding the correct room, which strengthens the statistical significance.

3.3. Trial Duration in Habituation, Common and Uncommon Trials

The mean duration of habituation trials is relatively high (33.69 seconds). This might be expected, as participants are initially unfamiliar with the task and take more time to complete the trials.

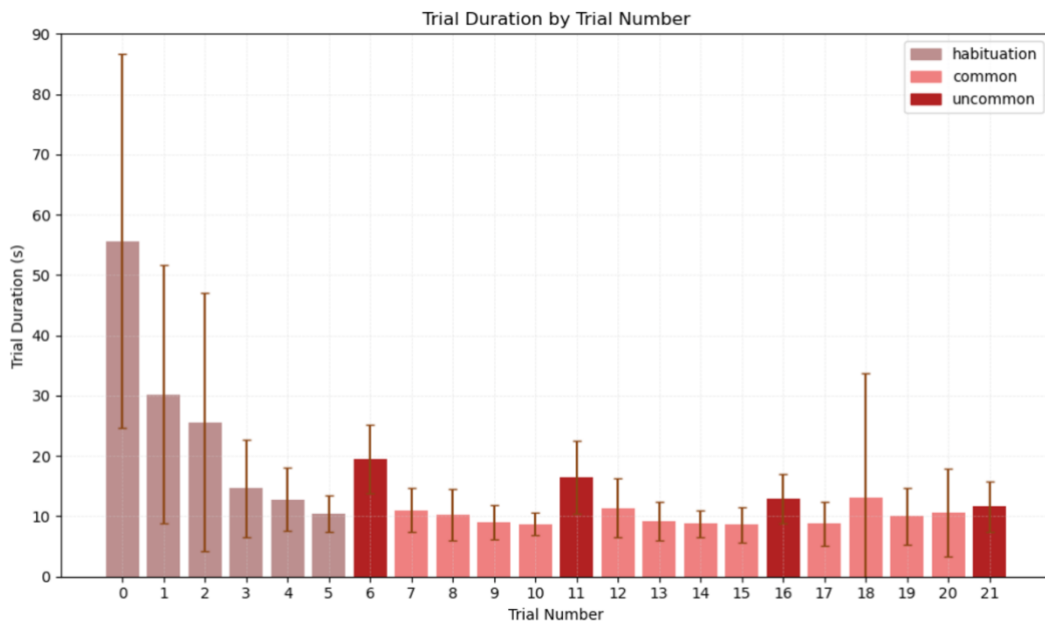


Figure 10: Trial Duration decrease in different phases

The mean duration of common trials is lower (11.0 seconds). This indicates that participants are becoming more efficient as they become accustomed to the task through habituation. As expected, the mean duration of uncommon trials is higher (15.96 seconds). These trials require participants to deliberately change their behavior, resulting in increased trial duration. This change could be due to participants having to think more or even making mistakes in these trials.

The error bars show the variability in trial durations within each trial group. Habituation trials exhibit higher variability, which might be due to participants' initial learning phase. Common trials have lower variability, indicating more consistent performance.

The graph effectively visualizes the relationship between trial numbers and trial durations for different trial groups. It demonstrates the trend of habituation leading to faster completion times over time. Uncommon trials deviate from this trend due to their deliberate nature, resulting in longer trial durations.

3.4. Corridor choice

The graph visually illustrates how participants (both Gamers and Non-Gamers) selected different corridor options in each trial.

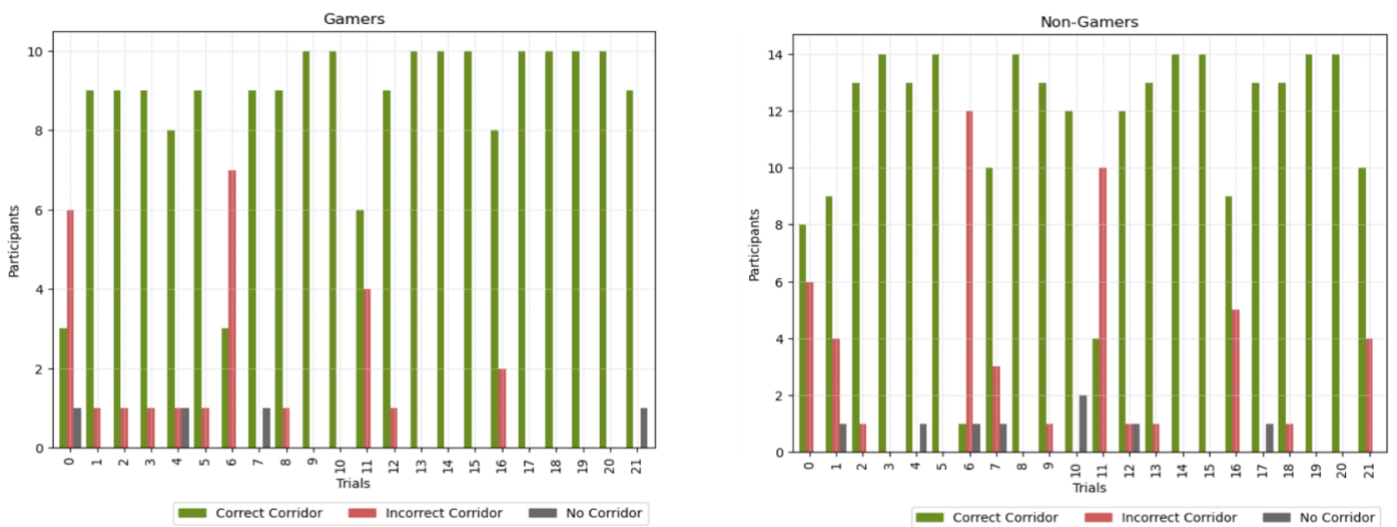


Figure 11: Corridor choice of each trial for Gamers and Non-Gamers;

In the first trial almost, all Gamers encountered errors when choosing the correct corridor, whereas Non-Gamers show a lower number, which is still not an indication of better performance but rather getting used to the game and exploring the environment. In the following trials after the first “test” trial, in the first 6 trials Gamers have the tendency to perform better after figuring out the correct corridor, only 1 participant seems to encounter errors in choosing the correct corridor. In the first uncommon trial, in trial 6, both Gamers, specifically 7 participants as well as Non-Gamers 12 participants make errors in navigating towards the correct corridor, which lies in the HER task anomaly, as predicted.

In the first re-habituation trial, which is trial 7 Gamers show no error in their corridor choice, Non-Gamers on the other hand, despite a low number of error, still 3 participants

have made an error in choosing the correct corridor. Non-Gamers 12 participants make errors in navigating towards the correct corridor, which lies in the HER task anomaly, as predicted.

Towards the ends of the other uncommon trials; 11,16 and 21 only four Gamers encounter errors in trial 11, then only 2 participants encounter errors and at the final uncommon trial no Gamers encounter any errors, remain to choose the correct corridor. Non-Gamers chose the incorrect corridor till the last trial, although significantly low number of participants did. In the case of “no corridor” participants chose as their first direction, the step to the elevator or tried to move backwards with the backwards key, since the experiment was not designed to let participants move backward but only up, left and right, this represents not a usual “error”.

3.5. Questionnaire

Among the pool of 26 participants, a small subset of 2 individuals had to be excluded due to being outliers. When solicited about the strategies they employed to rectify errors or missteps, participants revealed a recurring pattern. Several participants in the study exhibited distinct cognitive approaches when navigating the virtual environment. A prevalent technique involved the deliberate use of a specific mnemonic, wherein participants consciously encoded the elevator they had employed. This mnemonic served as a reference point for their navigation. Notably, among the Gamer subgroup, five out of the ten participants reported employing a calculation-based strategy.

Conversely, a subset of Gamer participants (three individuals) disclosed an employment of mental mapping techniques. These participants initially focused on establishing a sense of rotation to attain a comprehensive overview of the environment. By adopting this method, they aimed to enhance their spatial understanding before charting a path towards their goal. Remarkably, a subset of Gamer participants, distinct from the groups, highlighted a strategy centered on error awareness. They reported heightened attentiveness subsequent to an initial navigational error, thus enabling more cautious navigation.

In contrast, the Non-Gamer cohort predominantly exhibited a strategy characterized by orientation and limited rotation. Many participants within this group stated a lack of defined tactics but emphasized attempts to orient themselves within the environment. They noted a conscious avoidance of excessive rotation to prevent confusion arising from the environment's inherent symmetry.

4. Discussion

4.1. Trial Duration and Habituation

We adopted trial duration as a comprehensive indicator of the underlying decision-making process. This is grounded in the notion that as participants become accustomed to the task (habituation), their behavior tends to become more predictable and automated, leading to quicker trial completions. Conversely, when deliberation plays a role, uncertainty and hesitation manifest as prolonged trial durations. This underlying assumption is validated by the empirical findings of the study.

Notably, the measurement of trial duration in this task underscores the capacity of humans as rapid learners when faced with a limited number of trials. This is evident from the observable reduction in atypical trial behaviors in subsequent infrequent trials.

In essence, our study interprets trial duration as a meaningful reflection of the decision-making processes in action. The observed correlation between habituation, trial speed, and automation, as well as the reduction of aberrant behaviors in later uncommon trials, highlights the remarkable rapid learning ability of Gamers, in particular, as described in literature, due to their visuospatial skills, that enable to learn faster.

4.2. Habituation failure

Our hypothesis postulated that individuals who identify as Gamers would exhibit superior performance in locating the correct room during uncommon (6,11,16,21) and re-habituated trials (7,12,17) characterized by deliberation or habituation failure, as well as during the subsequent re-habituation phase. This proposition is aligned with the findings of Neto (2022), who elucidated that habituation failures arise when individuals encounter unforeseen obstacles or modifications in their environment after adapting to it over a specific duration. In the case of both Gamers and Non-Gamers, a similar principle holds true; however, Gamers tend to excel in effectively managing these infrequent and challenging situations, resulting in quicker responses.

Analyzing the data further, particularly observing the trajectory of trial durations towards the culmination of the experiment, we can discern intriguing patterns. The duration of time taken for the final uncommon trial (depicted in Figure 8) significantly diminished in comparison to the initial uncommon trial. This dynamic underscores the concept of the speed-accuracy trade-off, a principle widely discussed in the literature.¹⁵ Notably, this trade-off is particularly pronounced among Gamers. Their heightened familiarity with challenges, garnered from their gaming experience, translates into refined accuracy even as they navigate at accelerated speeds. This phenomenon is shown by their ability to swiftly adapt to the unique challenges posed by uncommon trials, as evidenced by their improved navigation times and enhanced accuracy.

In essence, this gives insights to support the notion that Gamers possess an advantage in managing the intricacies of uncommon trials, both in terms of faster responses and accurate decision-making. This advantage culminates in the acceleration of navigation times, particularly evident in the closing stages of the experiment.

4.3. Corridor choice

The exploration of participant behavior during the initial trials provides valuable insights into the learning dynamics of both Gamers and Non-Gamers within the experimental framework. In the inaugural trial, a substantial number of Gamers encountered difficulties when selecting the correct corridor, in contrast to Non-Gamers who exhibited a comparatively lower frequency of such errors. This disparity, however, is not indicative of superior performance among Non-Gamers but rather highlights the initial process of acclimatization and environmental exploration in the absence of prior gaming experience.

As the trials progressed beyond the initial "test" trial, a noteworthy trend emerged among Gamers. During the first six trials, Gamers displayed a propensity to improve their performance after discerning the correct corridor, with only one participant encountering errors in corridor choice. However, in the first uncommon trial (trial 6), a surprising occurrence unfolded. Both Gamers and Non-Gamers exhibited errors in navigating towards the correct corridor, as anticipated. This anomaly corresponds to the HER task prediction and underscores the challenge introduced by the unconventional trial.

Subsequent re-habituation trials (e.g., trial 7) revealed intriguing patterns. Gamers demonstrated a notable absence of corridor choice errors, indicating their adaptability to the re-established conditions. Non-Gamers, while exhibiting a lower but still notable rate of corridor selection errors (3 participants), displayed a consistent alignment with the HER task's anticipated anomaly in corridor choice errors.

Towards the concluding uncommon trials (trials 11, 16, and 21), an increasingly refined trend among Gamers becomes evident. Only four Gamers encountered errors in trial 11, which further dwindled to two participants in the subsequent trial. Impressively, in the final uncommon trial, Gamers demonstrated a complete absence of errors in corridor choice, indicative of their adeptness in navigating towards the correct corridor. In contrast, Non-Gamers continued to demonstrate corridor choice errors, albeit with a significantly lower number of participants.

Of particular interest is the behavior observed when participants selected a direction not corresponding to a corridor. Participants either opted for the elevator as their initial direction or attempted to move backwards, even though the experiment was confined to forward, left, and right movements. This divergence from conventional "errors" underscores the unconventional challenges presented by the experiment's design.

4.4. Limitations and Future Outlook

To dissect the habituation and deliberation mechanisms intrinsic to gaming experience, the experimental design focused on a 2D grid world interface exclusively operated using keys. This approach was motivated by the predominant nature of gaming experiences, often rooted in computer or console video games where key inputs drive interactions. However, this design choice carries certain limitations, particularly its simplicity and monotony. The experiment's structure was streamlined and repetitive, potentially compromising the depth of insights attainable.

To attain a more profound understanding of decision-making models in the context of gaming, a shift towards virtual reality (VR) could offer advantages. VR aligns better with the immersive nature of gaming, enhancing ecological validity. It provides a platform to replicate real-world scenarios, thereby potentially capturing the intricacies of decision-

making processes more authentically. By transitioning to VR, researchers can mitigate monotony concerns and create a more engaging environment that better mirrors the complexity of gaming experiences.

Another enhanced exploration of habituation and deliberation within the HER task could be facilitated by incorporating neuroimaging techniques such as fMRI or EEG. These methodologies could help bridge the gap by correlating established behavioral markers of deliberation and habituation with brain activity. Specifically, linking behavioral indicators associated with deliberation to hippocampal activity and those linked with habituation to dorsolateral striatal activity, as established in the literature, could provide valuable insights into the human decision-making process.

The issue of interpersonal variability in decision-making holds significance, yet has remained relatively unexplored. Understanding the influence of individual differences on decision-making results and behavior could yield valuable insights into maladaptive behaviors.

Consideration of the potential impact of participants' affect states on their performance, as well as on the interplay between habitual and deliberate systems, is a relevant avenue for future exploration. Incorporating ECG and skin conductance to measure decision making, especially habituation and deliberation under affective states. For instance, investigating whether factors such as fatigue, stress, sadness or distractions outside the experiment influence the preference for habitual or deliberate behaviors could provide a deeper understanding of decision-making dynamics. Additionally, delving into failures in arbitration between habitual and deliberate behavior, which have been associated with psychological conditions

In a broader context, the refinement of the grid world experiment can yield more insightful and nuanced results by implementing a series of enhancements. Diversifying the key assignments for participants stands as an initial step in enhancing the experiment's design. Introducing additional key options, such as the letters W, A, and S, not only minimizes potential confounding variables arising from keypress behaviors but also enriches the decision-making process, allowing participants to select from a wider array of directional choices. This alteration could potentially reveal more intricate patterns in decision-making strategies.

Another noteworthy augmentation involves the incorporation of a false door trigger sign. This ingenious addition contributes to a more sophisticated understanding of participants' behavior. By signaling when participants interact with or attempt to enter the false door, researchers gain a valuable tool for analysis. Examining how participants respond to this erroneous trigger provides insights into their decision-making adaptability and error recognition, potentially uncovering distinct cognitive processes during deliberation.

Expanding upon these enhancements, the introduction of a time limit emerges as a compelling avenue for investigation. By imposing a finite duration for task completion, the experiment delves into the interplay between decision-making and stress-inducing scenarios. The addition of time pressure can shed light on participants' abilities to navigate complex environments efficiently while maintaining accuracy. This dynamic introduces a heightened sense of urgency, mimicking real-world situations that demand rapid yet effective decision-making. Analyzing how participants prioritize and strategize under time constraints contributes a comprehensive layer to the understanding of their cognitive processes.

Furthermore, integrating visual and auditory cues could contribute to a more immersive and engaging environment. Visual cues could include changes in lighting, color, or object behavior, while auditory cues could range from subtle environmental sounds to more explicit feedback upon keypresses. These sensory cues provide participants with real-time feedback and further contribute to the overall experience and decision-making dynamics.

In summary, by expanding the key options, implementing a false door trigger sign, imposing time limits, and incorporating sensory cues, the refined grid world experiment not only minimizes confounding variables but also deepens our understanding of decision-making strategies, adaptability, and stress responses. These refinements align the experiment more closely with real-world decision-making scenarios and unlock a richer spectrum of insights for analysis and interpretation.

4.5. Conclusion

The HER task, modeled after real-world scenarios that induce habituation failures, serves as a reinforcement learning path-finding experiment. This task holds relevance to a broad spectrum of individuals due to its relatability to real-life situations. Within the scope of this master thesis, the principal objective is to conduct preliminary research into the efficacy of the HER task, now implemented within a grid world experiment, to investigate the mechanisms underlying the arbitration between habitual and deliberate behaviors in the realm of decision-making within video game players and non-video game players-

Our working hypothesis posits that the HER task, as adapted to the grid world, holds the potential to serve as an ecologically valid method for studying the behavioral manifestations of the arbitration process within gamers. We anticipated that relevant behavioral indicators of different modes of control would emerge from this experimentation.

The outcomes of our study align with existing literature, bolstering the concept that habitual behaviors manifest through repetition and knowledge of task sequences, resulting in shorter trial durations. In contrast, deliberation manifests as indecision, translating into longer trial durations. The collected data adds to the body of knowledge portraying humans as adept at learning quickly from a limited number of exposures, a phenomenon commonly referred to as "few-shot learning."

Though further investigations into the HER task's effectiveness are warranted, the initial findings validate its pertinence and applicability. The HER task, within the framework of the grid world experiment, simulates related habituation failures and enables the exploration of diverse decision-making modes. This endeavor has the potential to enhance our comprehension of both normal and cognitive skills and decision-making processes in the field of gaming, ultimately contributing to the development of interventions for future education, better understanding of gamers cognition and decision making, that might help to understand other neurophysiological fields.

Prospective research avenues could involve comparing behavioral correlates of habituation and deliberation with brain activity at decision points. Moreover, exploring interpersonal variability in decision-making could provide insights into potential associations with education, offering a holistic perspective on decision-making dynamics.

5. Bibliography

1. Gallivan, J.P., Chapman, C.S., Wolpert, D.M., Flanagan, J.R. "Natural Review of Neuroscience" 19, 519 (2018).
2. Siegel, M., Engel, A.K., Donner, T.H. *Front. Hum. Neurosci.* 5, 21 (2011).
3. Anguera, J.A., Boccanfuso, J., Rintoul, J.L., Al-Hashimi, O., Faraji, F., Janowich, J., Kong, E., Larraburo, Y., Rolle, C., Johnston, E., Gazzaley, A. *Nature* 501, 97 (2013).
4. Basak, C., Boot, W.R., Voss, M.W., Kramer, A.F. *Psychol. Aging* 23, 765 (2008).
5. Green, C.S., Bavelier, D. *Curr. Opin. Behav. Sci.* 4, 103 (2015).
6. Oei, A.C., Patterson, M.D. *Comput. Hum. Behav.* 37, 216 (2014).
7. Shiffrin, R. M., & Schneider, W. Controlled and automatic human information processing: II Perceptual learning, automatic attending and a general theory. *Psychological Review.* 84, 127– 190 (1977).
8. Lieberman, M. D. Reflective and reflexive judgment processes: A social cognitive neuroscience approach. In J. P. Forgas, K. R. Williams, & W. von Hippel (Eds.), *Social judgments: Implicit and explicit processes*. New York: Cambridge University Press. 44 – 67 (2003).
9. Sánchez-Fibla, M. Habituation and goal-directed arbitration mechanisms and failures under partial observability. *bioRxiv* (2020)
doi:10.1101/2020.11.24.396630.
10. Kahneman, D., & Frederick, S. Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics & biases: The psychology of intuitive judgment*. New York,
11. Loewenstein, G., O'Donoghue, T., & Bhatia, S. Modeling the interplay between affect and deliberation. *Decision.* 2(2), 55–81 (2015).
12. D. Kahneman. *Thinking, fast and slow*. New York: Farrar, Straus and Giroux, (2011).

13. E. Vertes and M. Sahani. A neurally plausible model learns successor representations in partially observable environments. (2019).
14. L. F. Neto and M. Sànchez-Fibla. Master thesis on Cognitive Systems and Interactive Media Behavioral Correlates of Habituation and Deliberation A virtual reality study with the Hotel Elevator Rows task. bioRxiv (2022).
15. Keramati, M., Dezfouli, A. & Piray, P. Speed/Accuracy Trade-Off between the Habitual and the Goal-Directed Processes. *PLoS Comput Biol* 7, e1002055 (2011).
16. R. S. Sutton and A. G. Barto. Reinforcement learning: an introduction, p. 526. isbn: 9780262039246
17. Naho Ichikawa, Greg J. Siegle, Alexandre Dombrovski, Hideki Ohira, Subjective and model-estimated reward prediction: Association with the feedback-related negativity (FRN) and reward prediction error in a reinforcement learning task, *International Journal of Psychophysiology*. Volume 78, Issue 3, (2010).
18. Collins, A. G. E. & Cockburn, J. Beyond dichotomies in reinforcement learning. *Nat Rev Neurosci* 21, 576–586 (2020).
19. Dayan, P. Improving Generalisation for Temporal Difference Learning: The Successor Representation. *Neural Computation* 5, 613–624 (1993).
20. Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J. & Daw, N. D. Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS Comput Biol* 13, e1005768 (2017).
21. Gershman, S. J. & Daw, N. D. Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annu. Rev. Psychol* 68, 101–128 (2017).
22. Momennejad, I. et al. The successor representation in human reinforcement learning. *Nature Human Behaviour* 2017 1:9 1, 680–692 (2017).
23. Edward C. Tolman. Classics in the History of Psychology -- Tolman (1948). *The Psychological Review* , 55(4) 189–208 (1948).

24. Stachenfeld, K. L., Botvinick, M. M. & Gershman, S. J. The hippocampus as a predictive map. *Nature Neuroscience* 2017 20:11 20, 1643–1653 (2017).
25. Jordan, T., & Dhamala, M. Video game players have improved decision-making abilities and enhanced brain activities. *Neuroimage: Reports* 2, 3 (2022).
26. Green, S., & Bavalier, D. Action-video game experience alters the spatial resolution of vision. *Psychol. Sci.* 18, 88-94 (2007).
27. Smith, E.T., Basak, C. A game-factors approach to cognitive benefits from video-game training: A meta-analysis. *PLoS One* 18, e123456 (2023).
28. Hamlen, K. Stochastic frontier estimation of efficient learning in video games. *Comput. Educ.* 58, 534-541 (2012).
29. Strobach, T., Frensch, P., & Schubert, T. Video game practice optimizes executive control skills in dual-task and task switching situations. *Acta Psychol.* 140, 13-24 (2012).
30. Moran, R., Keramati, M., Dayan, P. & Dolan, R. J. Retrospective model-based inference guides model-free credit assignment. *Nature Communications* 2019 10:1 10, 1–14 (2019).
31. Vértes, E. & Sahani, M. A neurally plausible model learns successor representations in partially observable environments. *Adv Neural Inf Process Syst* 32, (2019).
32. Redish, A. D. Vicarious trial and error. *Nat Rev Neurosci* 17, 147 (2016).
33. Hasz, B. M. & Redish, A. D. Deliberation and procedural automation on a two-step task for rats. *Front Integr Neurosci* 12, 30 (2018).
34. Santos-Pata, D. & Verschure, P. F. M. J. Human vicarious trial and error is predictive of spatial navigation performance. *Front Behav Neurosci* 12, 237 (2018).
35. Huynh, T., Alstatt, K., Abram, S. v. & Schmitzer-Torbert, N. Vicarious Trial-and- Error Is Enhanced During Deliberation in Human Virtual Navigation in a Translational Foraging Task. *Front Behav Neurosci* 15, 43 (2021).

6. Appendices

6.1. Questionnaire



HER Experiment

Questionnaire

Please fill out this quick questionnaire about your general experience with the HER environment.

Participant ID:

Sex:

Female Male Prefer not to say Other:

Age:

Previous experience with gaming environments?

- Very experienced
- Somewhat experienced
- Neither experienced nor inexperienced
- Somewhat inexperienced
- Very inexperienced

Did you experience errors or mistakes of any kind? Please, describe them.

Did you apply a strategy to solve your errors or mistakes?

How many mistakes do you think occurred during the experiment?

6.2. Notebook for Data Analysis

Notebook for data analysis, can be found under following link:

<https://github.com/akc7774/DataAnalysis.git>