Engineering Social Learning Mechanisms for Multi-Agent Interaction

by

Owais Hamid

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2021

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

Supervisor(s): Kerstin Dautenhahn
Professor, Electrical and Computer Engineering,
Systems Design Engineering, University of Waterloo
Chrystopher Nehaniv
Professor, Systems Design Engineering,
Electrical and Computer Engineering, University of Waterloo

Internal Member: Mark Hancock
Professor, Management Sciences, University of Waterloo

Internal-External Member: Kate Larson
Professor, Cheriton School of Computer Science, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners. I understand that my thesis may be made electronically available to the public.

Statement of Contributions

The thesis includes first-authored publications in the form of one accepted and published conference proceeding, published by the Institute of Electrical and Electronics Engineers (IEEE) and one that is currently under review by the Association for Computing Machinery (ACM).

Chapter 3 presents the first study that was designed by Owais Hamid with major contributions from a design perspective by Professors Kerstin Dautenhahn and Chrystopher Nehaniv (the supervisors). The experiments were performed on the Robotarium (a remotely accessible Robotics platform at the Georgia Institute of Technology - Atlanta, GA) and the manuscript was written by Owais Hamid, with corrections and recommendations from the supervisors.

Citation: O. Hamid, K. Dautenhahn and C. L. Nehaniv, "Engineering Social Learning Mechanisms for Minimalistic Multi-agent Robots," 2020 3rd International Conference on Control and Robots (ICCR), 2020, pp. 90-99, DOI: 10.1109/ICCR51572.2020.9344158.

Chapter 4 presents the second study. This study was designed by Owais Hamid with major help from Dr. Shruti Chandra, and recommendations provided by the supervisors further fine-tuned the project. The Unity game implementation and experimental procedures were implemented by Owais Hamid, and the questionnaires used in the study were designed on Qualtrics by Shruti Chandra. Manuscript preparation was done mostly by Owais Hamid with input and feedback from the other authors. The manuscript is under consideration for the Ninth International Conference on Human-Agent Interaction (HAI 2021).

Abstract

This thesis is strongly inspired by literature on animal social learning, applying it to multi-robot as well as human-robot interaction scenarios, Social learning, which can include complex or simple social mechanisms, allow us to understand cooperation and communication in animals, giving them better chances to survive for longer and thrive as a society. For this dissertation, to translate this understanding into socially rich behavior among multi-agent robots and Human-Robot Interaction, two experiments were conducted.

The first experiment focused on how social learning might optimize cooperation among robots (in a robot 'society') for the problem of foraging. The task utilizes small and simple swarm robots to understand how such social mechanisms might play a role in establishing rules for emergent group behavior and how social rules might be engineered to gain useful effects in a group of robots. The study investigated exploratory behavior without interaction (asocial) and with interaction (social). The results from this exploratory study suggest that deterministic asocial exploration is best performed by a Spiral exploration mechanisms. However, these asocial exploration strategies are eclipsed by certain types of social reward sharing strategies as long as sharing occurs for at least half the lifetime of the robots. Sharing locations of reward caches for all time is of course the most optimal, but comes at the cost of communicating longer and hence using more energy both on the sender and receiver's end. An analysis of a compromise strategy between completely asocial exploration and social reward location sharing is performed using strategies termed critical and conditional learning. It is found that the number of reward caches located through critical and conditional learning are intermediary to the two extremes, namely completely asocial and completely social foraging.

The second experiment sought to understand if and how other types of social learning mechanisms such as observational conditioning can facilitate social information spread to human participants. The question of whether, and to what extent, a robot can influence a human's actions is asked through a study designed to understand if emotions displayed by a robot demonstrators can influence human observers. An immersive first-person gaming experience utilizing Unity was designed where a robot demonstrator reacted either positively or negatively to an external stimulus. Objective (position of player in-game) and subjective (Questionnaire) data collected on the human participants' reactions suggests that the virtual robot agent is successful in socially transmitting information.

Through these studies, I seek to contribute to the understanding of the role simple social learning mechanisms can play in information transfer among human and robot agents, and to identify useful metrics for the detection of such social mechanisms.

Acknowledgements

I would like to acknowledge the contributions both my supervisors have made to my understanding of the world of robotics. My objective has always been to establish myself as an independent thinker with a deep understanding of Intelligent Systems, and training for this was provided to me in unique and colorful ways not just through my course work, but also through the rigorous research training both Professors Kerstin Dautenhahn and Chrystopher Nehaniv put me through, and I will forever be grateful for this. As an intermediary, someone who is not a Professor but someone who also has considerably more experience than me, I had the honor of working with Dr. Shruti Chandra, one of the most frank and professional Post-Docs I have ever worked with and I am grateful to her for keeping me focused and on point with respect to my experiments.

I would like to extend my gratitude to Professors Kate Larson and Mark Hancock for providing valuable feedback on the various aspects of Multi-agent Robotics and Socially Intelligent Robotics that this thesis touches on.

I have made some great friends within the Socially Intelligent Robotics Research Lab (SIRRL), and I would specifically like to thank Aishwarya Aravamuthan, Pourya Aliasghari, Hamza Mahdi, John Muñoz and Katrin Fischer for the amazing discussions we have had over the course of my studies. I would also like to thank Patrick Shaghaghi, Moojan Ghafurian and Mahsa Golchoubian for their help with some of my projects. Patrick has been an absolute delight to work with and I appreciate the different perspectives he brought into my work.

Lastly, I am grateful to the variety of friends I made through various other programs that ran on campus at the University of Waterloo. This provided me with a busy social life, even if it was overshadowed by the pandemic.

This research was undertaken, in part, thanks to funding from the Canada 150 Research Chairs Program and the participants who gave their time to the Experiments.

Dedication

I dedicate this thesis to my parents, especially my father who taught me to never give up, and to S, who taught me that it doesn't all have to be black and white.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

> It is not knowledge, but the act of learning, not possession but the act of getting there, which grants the greatest enjoyment.
>
> — Carl Friedrich Gauss, Letter to Farkas Bolyai (2 September 1808)

The spread of social information and the concept of learning from others, either through demonstration or some other means, has been the subject of study in several research fields for a variety of reasons. The most important advantages of social learning include adaptability through learning from others, which enables a variety of humans and other animals to increase their chances of survival and for them to thrive.

In a similar fashion, earlier work in robotics has argued that robots capable of learning from either humans or other robots can utilize information obtained elsewhere for their own benefits, and can therefore allow them to be robust to changes in the environment [1]. Social information might come from other robots or humans, and for robots to be effective, social information from both other robots and from humans needs to be extracted.

1.1 Problem Definition

Several mechanisms have been identified as playing a major role in everyday life. Some of these are cognitively more complex and require greater cognitive capacity. Examples of such social transmission mechanisms include various forms of imitation. Other forms of social transmission might involve methodologies as simple as redirection of attention, such as Local Enhancement.

While we have discussed the potential benefits of social learning among robots or between robots and humans, in nature social learning is not always beneficial [37]. Information that is passed between agents might be outdated if the environment has changed. Further, the forms of social learning utilized among biological species may not be suitable for information propagation among robots. There is, therefore, a need to further investigate whether such forms of social learning might actually benefit robots and under what conditions these benefits apply.

Within the forms of social learning, a wide range of social transmission mechanisms might be possible, ranging from imitation on one end of the complexity spectrum to local enhancement as the simplest on the other end, with many others in between. While imitation happens to be the most studied form of social learning in Robotics (an excellent review on what is termed "Learning/Programming from Demonstration", or just "imitation" is found in [61]), challenges based on cognition and perception of an agent, and defining how close an action is to what has been observed remains an unresolved challenge. A bigger problem seems to be the fact that despite how humanoid a robot might be, it still does not have the same *affordances* that a human might, neither does it have the exact same embodiment, which means the agent needs to map the actions it observes on to its own embodiment. This can be summarized as the correspondence problem [54], and remains an active area of research.

This study, therefore, seeks to avoid the complexity of imitation and explore metrics of observation-based learning between multiple robotic agents, and between robots and humans that are *simpler* (than imitation) in nature.

1.2 Objectives and Contributions

The work incorporated in this thesis draws inspiration from and seeks to replicate certain experiments based on behavioral sciences with two broad objectives in mind: establishing a proof of concept of simple forms of social transmission taking place between a robot and (a)

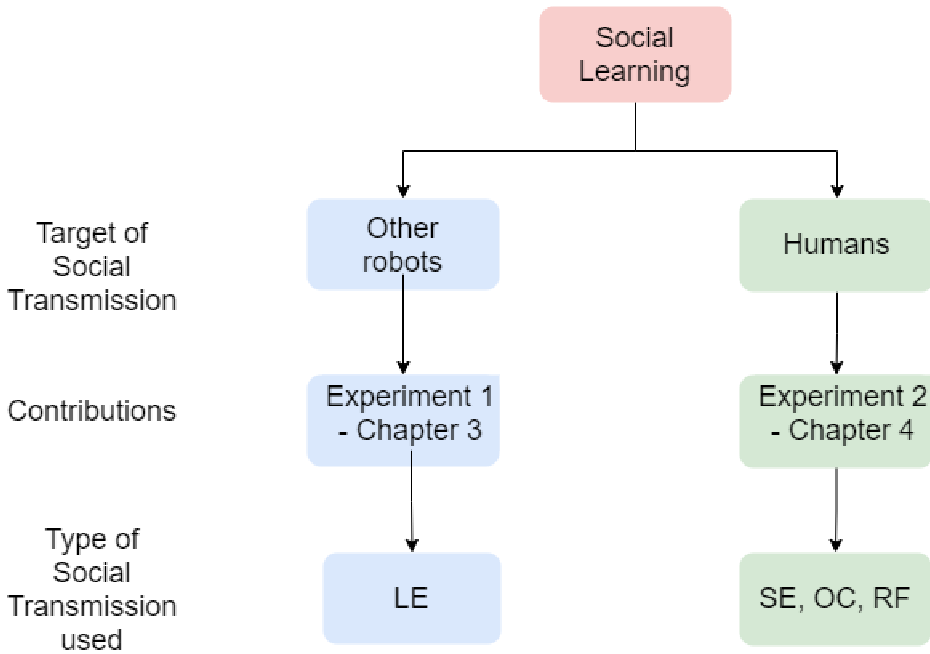


Figure 1.1: A summary of experiments conducted during the dissertation, with contributions and types of social transmission used. SE - Stimulus Enhancement, LE - Local Enhancement, OC - Observational Conditioning, RF - Response Facilitation.

another robot, in which case this becomes relevant to Multi-agent Robotics, or (b) another human, in which case Human-Robot Interaction takes place. A flowchart that describes the methodology that this thesis follows is portrayed in Figure 1.1. Here, SE stands for Stimulus Enhancement, LE - Local Enhancement, OC - Observational Conditioning, and RF - Response Facilitation. It is to be noted that Multi-agent robotic interaction as specified in the blue branch was performed through Experiment 1, which is explained in Chapter 3 and transmission of social information between a robot and human participant shown in the green branch (Human-Robot Interaction) happens in Experiment 2 which is explained in Chapter 4.

Certain specific types of social mechanisms of information transfer have been utilized to perform Multi-robot interaction for both experiments. Local enhancement has been chosen as the subject of the first study because it is one of the simplest forms of social learning that is found in animals [70], the idea being that some of the simplest forms of social learning may enhance the capabilities of robot learning, and make learning computationally cheaper than the complex deep learning algorithms for imitation proposed elsewhere ([77][23][45]). Further more, it is difficult to find cross-disciplinary studies across robotics and psychology that focus on the subtler, less known, but also computationally cheaper forms of social learning. Thus, we focus on simpler algorithmic complexity for a minimal real time computation scheme.

For robots to influence human decision making and experience, clues can be taken from experiments performed in behavioral sciences. The inspiration for Experiment 2 comes from humans interacting with animals, other (less experienced) humans and social information transfer across species. For such processes to take place, the most likely candidates of social transmission mechanisms are Stimulus Enhancement, Observational Conditioning and a few others. All these are defined and examples of their use, along with explanations of their relevance to the current experiment are explained in Chapters 2 and 4.

1.3 Overview of Thesis

The thesis is divided into four further chapters (making for a total of five, including this chapter). A literature review that touches on topics that are utilized for the two experiments is performed in Chapter 2. This chapter goes over concepts from topics such as social learning among animal/primate species, between different animal species and humans and among humans, along with previous work utilizing social learning in robotics. In Chapter 3, a specific Social Learning mechanism is utilized for robot-robot interaction constituting a Multi-agent Robotics experiment. A different set of social learning mechanisms are utilized for transmission of information between a robot and a human in a Human-Robot Interaction experiment in Chapter 4. Observations regarding the two experiments and how this advances our knowledge of Socially Intelligent Robotic systems is made in the concluding chapter, Chapter 5.

Appendix A further includes the ethics certification utilized for the study described in Chapter 4. Appendix B further describes the game experience for the same experiment, and all the questionnaires administered. A very basic python-based algorithm that replicates some features of the Multi-agent Robotics experiment described in Chapter 3 is also described in Appendix C.

Chapter 2

Literature Review

> Research is formalized curiosity. It is poking and prodding with a purpose.
>
> — Zora Neale Hurston, 1942

Natural biological systems interact and continuously exchange social information through a variety of mechanisms. These mechanisms of exchange are collectively classified under social learning, which seeks to understand how observation of each other can cause agents to perform the same body movements, go to the same place as the other agent went in order to investigate and discover, utilize an object for similar purposes, make similar sounds, or think similar thoughts [56].

The advantages that social learning presents for animals or humans are varied and will be discussed further later on. There is however strong literature based support which proposes that machines or robots capable of taking advantage of social learning, given the unstructured and complex data provided by social context, can take advantage of lessons learned earlier (by humans or other robots) and hence will be much more robust to changes in environment and be better able to survive[22].

The first part of this review seeks to explore Social Learning in animals and humans. This includes the various classifications of social learning that occur in nature. The second part wishes to understand previous work on social learning that has been applied in robotics.

2.1 Social Learning in Nature

2.1.1 What is Social learning?

By definition, social learning is defined as “learning that is influenced by observation of or interaction with another individual or its products” [41]. This is of course, contrasted with asocial or individual learning which doesn’t require as many factors to be accounted for. Asocial learning becomes very expensive in cases that involve avoiding predators, food (whether its poisonous or not) and other stimuli. Nevertheless, social learning might incorporate outdated practices which might be detrimental in a changing environment with threats such as predators that evolve.

Hoppitt and Laland re-define Social Learning slightly differently[43]

> Social Learning is learning that is facilitated by observation of, or interaction with, another individual (or its products)

In this section, the focus of the review is on the types of social learning mechanisms, along with how these mechanisms are observed in nature in animal taxa.
It is first of all important to define the roles in social learning. The agent learning a task is henceforth called an **observer**, whereas the agent portraying a certain behavior is termed **demonstrator**.
Galef [7] defined *Social Transmission* as “cases of social learning that result in increased homogeneity of behavior of interactants that extends beyond their period of interaction”. Social transmission is mentioned here to try and narrow down the definition of social learning and contextualize what happens when scientists observe animals displaying ’intelligent’ behavior through social learning. It is different from Imitation, where the definition of imitation, to be formalized later, means something different in the context of Social Learning.

Social transmission has been given a more formal/logical definition by Hoppitt and Laland [43]:

> Social transmission occurs when the prior acquisition of a behavioral trait T by an agent A, when expressed either directly in the performance of T or in some other behavior associated with T, exerts a lasting positive causal influence on the rate at which another individual B acquires and or performs T.

Social transmission is further defined into Social transmission of trait acquisition, which defines how long it takes for B to learn T, and social transmission of trait performance, which defines the rate at which B is able to perform T.
Since Social learning is the subject of our study, and as can be clearly seen, most of this learning is shared across generations, it is perhaps also useful to define both tradition and culture. Fragaszy defined tradition as "a distinctive behavior pattern shared by two or more individuals in a social unit which persists over time and that new practitioners acquire in part through socially aided learning"[26].
Culture, while being a very tricky subject and immune to a commonly agreed definition, is defined by Hoppitt and Laland as"those group-typical behavior patterns shared by members of a community that rely on socially learned and transmitted information" [43].

2.1.2 Why Study Social Learning?

It is important to understand why the study of Social Learning is essential.

Firstly, imitation is considered one of the primary forms of transmission of information, and it is important to understand how this process happens. One of the most pressing problems that has been identified, in both natural social learning and in engineering (mathematical) approaches[55] for imitation has been the correspondence problem. The problem can be summarized by the question, how does the brain convert the perception of an observed act into an action performed by one's own embodiment? [43]. A good example of the correspondence problem is figuring out where the cats went from time 1 to time 2 as shown in Figure 2.1. One cat could be larger and the other smaller, as shown in the upper figure and so we can assume they followed a linear distance. But the smaller cat might also be small because of distance (from our perception) and might have moved closer hence becoming larger, with the opposite happening to the erstwhile larger looking cat.
The consequences of the discovery of mirror neurons (neurons that are active during observation and application of the same actions as seen in someone else) for the correspondence problem have been debated, with some, such as [17] and [40], that call them mechanisms that enable imitation, or perhaps a by-product of social learning. This is of prime importance to socially intelligent robotics because, as we shall see, imitation overlaps with several lower forms of social learning.

Secondly, various types of social learning, including imitation, are displayed by animals including insects, and it seems to provide them with making adaptive decisions about their environment, and this allows them to forage more efficiently and avoid predators, contributing to their survival [36]. This, and the significant effect of imitation and other

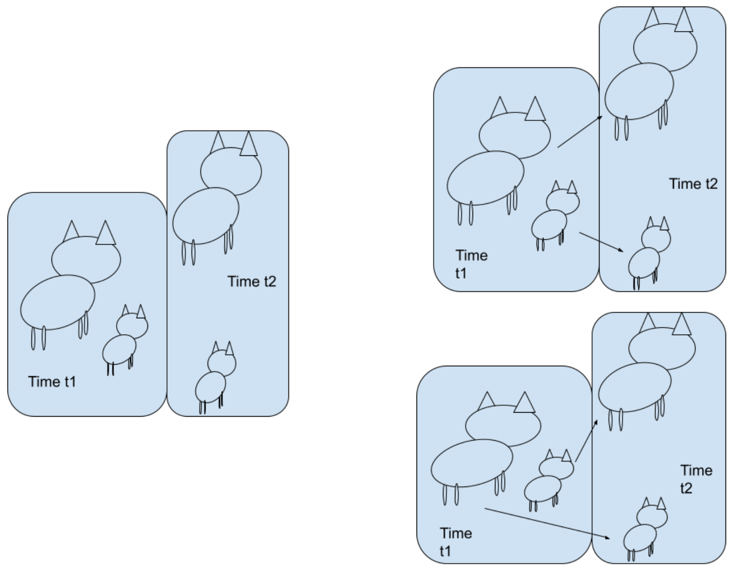


Figure 2.1: An example of the correspondence problem. Inspired by [67]

forms of social learning in the development of babies and infants that effect the social life (hence related to other disciplines such as sociology and developmental studies) of children seem to be strong motivators for study of social learning[43].

Thirdly, social learning within animals has given rise to theories regarding animal culture and traditions within so called 'societies' of various animal taxa such as in apes[74] or cetaceans[46]. [1] Our concern about societies and/or culture pertains to how robots can influence societies and make certain experiences better for humans.

However, central to this study of imitation or social learning is the question of how social learning can be scaffolded to make robots more intelligent. Of course, this involves going well beyond simple Artificial Intelligence, and requires inspiration from social cues in living beings, not just animals but in humans, especially as they develop from babies to grown adults.

[1] While it is exciting to talk of 'societies', 'culture' and 'tradition' in animals, there has been fierce criticism about whether those terms should be used in the context of animals. There are scientists who believe that animals show pseudo-imitative or semi-imitative phenomena[69][43] which are quite different to social learning, and are more instinctive in nature. Perhaps it is easier to classify them as proto-culture or pre-culture, as described by Goodall[38] and Whiten [75]. Tomasello[72], Galef[31], and Laland [48] are of the opinion that comparisons of culture are superficial and not grounded in cognitive processing. Definitions of social learning and imitation, culture and tradition depend on the context (biology, psychology etc.) and the looser the definition, the more inclusive it is[43] of other creatures.

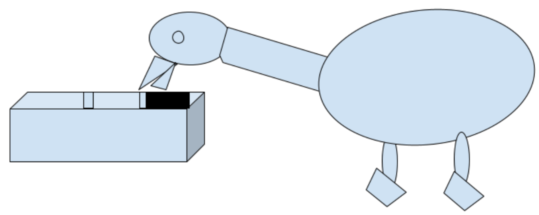


Figure 2.2: Greylag goose opening a box with its bill, taken from [28]

2.1.3 Social Learning Mechanisms

A number of efforts have been made to classify social mechanisms of information transfer in primates, and the definitions have changed over the past decades according to leading authorities. Some of the older efforts were made by Galef [32], Heyes [41] and Zentall [79]. The definitions have continued to evolve over time, and the ones utilized here are more recent versions built on previous work as defined by [43].

Stimulus Enhancement

Stimulus Enhancement as defined by Heyes [41] occurs "*when an observer observing a stimulus at time t1 from a demonstrator and this stimulus causes an observable change in the observer at time t2*".

A good example of Stimulus enhancement is recorded among the greylag geese (*Anser anser*) who, when observing humans open a box, are able to open the box themselves much quicker than those that haven't been exposed to humans opening a box[28]. The study with the greylag geese is further interesting because it shows an instance where a human provides social information to an animal. Further, because their embodiments are different, the human demonstrator showed how to open a wooden platform with a single finger, and the geese opened the platform with their beak/bill. The inspiration for this study came from the observation that the geese had started utilizing a different form of feeding from stems of butterbur (*Petasites hybridus*), where they bit through the stem and chewed exposed stems. The observation of an experienced model helped the observers learn themselves how to feed this way.

Opportunity providing

Opportunity providing, also called scaffolding, is when a demonstrator places the observer with an advantage by providing the means necessary to discover the 'correct' or most efficient way of solving a problem[43]. Examples of opportunity providing can be found among black rats (*Rattus rattus*) in the forests of Israel, who strip pine cones[2] if they have a mother or foster-mother that also strips the pine cones. Further, adult rats do not strip them even if they find evidence of this happening. A better example would be mother cats providing kittens with opportunities to learn and familiarize themselves with hunting rats [16]. Mothers bring slightly incapacitated rats which kittens then kill, providing them with the opportunity to develop their predatory skills.

Local Enhancement

Local Enhancement as a term was first defined by Thorpe[70] who defined it as the refocusing of attention of an animal to a specific object or part (location) in an environment. The attention of a 'naive' learning agent is attracted towards a specific location by a demonstrator agent, because this location might hold information that contributes to the task at hand[43].

Local enhancement was suggested to be a subset of stimulus enhancement since the stimulus being enhanced could just be the location. However, local enhancement may occur without learning, a detail of some importance, since such non-learning processes may then lead to social learning. A model may simply attract the observer to its location, and other agents might simply come together at that particular location.

The formal definition as given by [43] of local enhancement follows:

> Local Enhancement occurs when, after or during a demonstrator's presence, or interaction with objects at a particular location, an observer is more likely to visit or interact with objects in that location.

A good example of Local Enhancement is the experiment conducted by Reader et al. on female guppies (*Poecilia reticulata*). Experiments in the wild were conducted about feeding patterns, and it was concluded that one of the foraging sites was preferred by the guppies due to the presence (location enhancement) of conspecifics. A similar situation arose with guppies preferring to escape through a route demonstrated by their conspecifics when a trawler net simulated a trap the fish are well aware of[62].

More direct evidence of local enhancement is found in *Bombus terrestris*, Bumblebees [49]. Hoppitt and Laland [43] countenance this experiment to be proof that local enhancement and observational conditioning are not the same. The primary problem bumblebees face is collection of nectar as efficiently as possible. Sucrose rich yellow flowers (reward) and blue flowers without sucrose (failure) were presented to a group of bees, and the first group learned the correct flowers to collect from. It was found that when the sucrose rich yellow flowers were visited by demonstrators, a second group of observers overwhelmingly preferred yellow flowers. Moreover, it was not the characteristic of flowers being yellow that compelled the observers, rather it was the flowers that were visited by demonstrators that were also visited by observers, thus establishing local enhancement as opposed to stimulus enhancement as the social mechanism at play.

Imitation

Among the most difficult to define mechanisms of social transmission is Imitation. The difficulties arise from three main factors as defined below [43] :
A capacity for **Intentionality** is seen by some (see Tomasello [71]) as evidence of *intentional* copying, which means for copying to be constituted as imitation according to Tomasello, intentions must be observed. However, a weakness of this approach is that intentionality is not measurable [78], besides which there is no reason for such intentionality to be integral to imitation [10].
How **accurately** does the observer copy the demonstrator is another issue that has been raised by Nehaniv and Dautenhahn [21]. The accuracy in question is regarding how similar the actions, states and effects are of an observer with respect to the demonstrator's.
Lastly, **novelty** of actions is considered a necessary condition by certain authors, i.e. imitation cannot happen if the actions are not novel to the observer [13]. There is some disagreement here, and it is important to be careful regarding this point.
Novelty is considered the primary point of distinction among the above three classifiers that have been used to distinguish original imitation from lesser forms of imitation by both [11] and [43]. This helps them classify imitation into two types.
Contextual Imitation is defined by [43] as when an observer watches the demonstrator

perform an action in a specific context, and this makes them more likely to perform the same action in the same context. **Production Imitation** occurs when a demonstrator performs actions not in the observer's repertoire and the observer becomes more likely to reproduce the same sequence after observing the demonstrator. The focus here is on the novelty of the sequence of actions and whether direct observation is used to reproduce these actions.

Response facilitation

Response Facilitation (RF) is defined as the "presence of a demonstrator performing an act [that] increases the probability of an animal which saw it do the same" ([9] p.237). A good example is [12] where a western lowland gorilla, *Gorilla g. gorilla*) learns to reproduce observed actions from a human. The gorilla then puts together actions it already knows, and then reinforces them by individual learning and copying gestures that humans showed her. Naive coders, who are asked to check if such actions resembled the original gestures by humans, gave scores that confirmed that the actions taken by the gorilla were very similar to the human's. This study is another instance of animals extracting information socially from humans.

Observational Conditioning

Observational Conditioning traditionally refers to Pavlovian conditioning where an *Unconditioned Response (UR)* to a stimulus for the demonstrator acts as an Unconditioned Stimulus (US), which then becomes a *Conditioned Stimulus (CS)* for the observer. The observer then responds to the CS in the same way as the demonstrator did [43]. Heyes' definition is a little broader in that stimuli need not be conditioned [41]. It is defined as "*a subset of Stimulus-Stimulus learning where an observer observing a demonstrator is exposed to a relationship between stimuli at t1 and the observer's behavior changes in a detectable manner at a later time t2*".

Particularly important to us is a classic study performed by Gerull and Rapee [35] that serves as an inspiration for experiment 2. Here, they prove that affective responses to novel objects can be learned socially. Specifically, if humans observe something novel, they usually associate certain emotions to the newly introduced object of interest when they observe someone else who has more experience react to the novel object or situation in a certain way. In [35], toddlers are introduced to novel toy creatures and observe their mothers reacting to the objects with fear, horror or disgust and avoidance. This conditions them to associate fear and avoidance behavior towards the objects when they witnessed their mothers do the same. Toddlers reacted significantly differently when a mother's reaction is positive, not showing any avoidance to the toy physically. Facial expressions are also very different in the two cases, the children in the fear / disgust condition showing expressions of clear disgust or horror, while those in the positive condition treat the toy normally. This shows that the information provided by an experienced demonstrator (e.g. mother) is crucial to the perception of any object that the observer (e.g. toddler) observes, hence creating a Stimulus-Stimulus pairing. Gerull and Rapee also found significant differences between the perception and avoidance behavior of male toddlers and female toddlers to the mother expressing a positive or negative emotion. Male toddlers were less avoidant towards the toy creature in the negative condition than female toddlers. Furthermore, the impression of these emotions lasted for at least up to 10 minutes, making this one of the first studies to understand potentially how long the effects might last.

Differentiating and Detecting different types of Social Learning Mechanisms

Following from [43] and Figure 2.3, **Stimulus Enhancement** can be detected if it is shown that the social transmission of a trait exists. This can be proved by showing that the observer displays an increased response rate to the trait it has observed, and that this response to the stimulus increases the efficiency with which the trait is performed[43].
A good example is the choice of mates among female quail (*Coturnix japonica*). The females tend to choose males they observe around other females, hence copying the choices their peers make[33]. Local enhancement was ruled out because the choices don't limit female quails to a specific location. There are no clear Stimulus-Stimulus (S-S) association, and hence observational conditioning is ruled out as well[43].

Local Enhancement can be distinguished from other types if it is shown that if a demonstrator leads naive observers to a specific location, the acquisition of a trait happens at a faster rate than at other locations. The example of the *Bombus terrestris* given in section 2.1.3 provides the clearest evidence of local enhancement[43].

Observational Conditioning is detected when an observation of a stimulus creating a certain response in a demonstrator also creates the same response in the observer. Cook et al.[19] exposed a monkey who had fear of snakes to another observer monkey. The result, predictably, was the development of the same fear in the observer due to S-S conditioning.

An excellent methodology to separate the types of social learning that might be happening in a situation is to use the decision tree shown in Figure 2.3. Here, any type of social transmission that happens is divided into two major groups, Location specific or

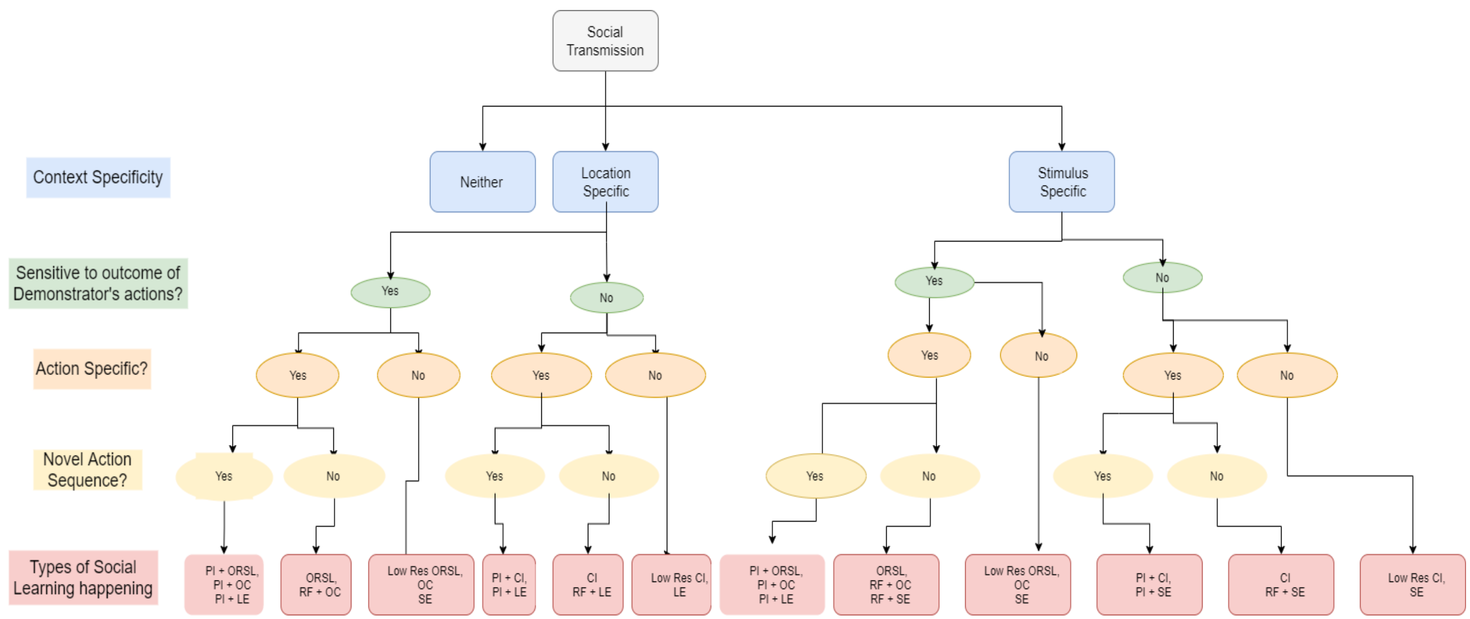


Figure 2.3: A decision tree describing which types of social transmission are at play depending on the effect of the stimulus and the demonstrator's actions, adapted from [43]. Here, PI = Product Imitation, CI = Contextual Imitation, ORSL = Observational R-S Learning, OC = Observational Conditioning, LE = Local Enhancement, RF = Response Facilitation, SE = Stimulus Enhancement.

Stimulus specific. This is further dependent on whether the social transmission is sensitive to the demonstrator's actions, if it is action specific and whether the action sequence is novel or not. Depending on this, the social transmission might be classified into the various categories as shown in the red boxes to the bottom of the Figure.

2.1.4 Strategic Social Learning

It would be overly simplistic to assume that copying or learning from others is to avoid the trial and error repetition, or perhaps due to minimizing energy expenditure (out of laziness). Giraldeau [37] considered social learning as information parasitism. This is because while social learning is more about information flow, asocial learning is more about finding new information. Since the environment is constantly changing, asocial learners need to be constantly sampling the environment. Should this not happen, learning will happen only over outdated information Since asocial exploration is what 'samples' the world and finds original information, social learning is only effective if the number of social learners is rare as compared to asocial explorers. Alan Rogers utilized this to propose what is now known as Rogers' paradox [64], that a social learner's ecological fitness (to survive an environment) could be at most that of an asocial explorer's. It was, however, pointed out that social learner's can learn from asocial explorers and then build on top of that knowledge, making them more fit to survive an environment [24].

Therefore, the question of **when to copy** needs to be addressed here as well, distinguishing between **Critical and Conditional** social learning. Critical learning happens when individuals learn socially first and then if the solution provides little value, the individual improves upon it to make it more efficient. The reverse, Conditional learning, happens when individuals attempt to learn asocially first, but because of certain failures move on to learning socially from other agents.
Critical social learning has been termed an Evolutionarily Stable Strategy (ESS) [47]. It is helpful when asocial exploration is either too costly or the solution is too complex for a solution to be discovered by a single agent. However, a highly variable environment would mean it is better to asocially explore the environment first and this is where conditional learning might be more beneficial.

The inspiration to copy could be widely varied and agents might be induced to copy because of reasons ranging from the established behavior being not useful to the common situation of when asocial learning is simply too costly. The costs associated might include energy expended, but in cases of anti-predator behavior, this could mean being eaten by the predator. A good example is the nine-spined stickleback fish (*Pungitius pungitius*).

Among these fish, the foraging preferences are almost exclusively based on observational learning i.e. learning foraging routes and feeding areas based on conspecifics. This is because in open water, these fish are prime prey for piscivorous (fish-eating) predators due to their inferior skeletal structure making them easier targets and therefore they rely exclusively on information passed on socially.

2.2 Social Learning in Robotics

A good understanding of the previous work that has been done on Social Learning in Robotics is essential for the studies that were designed in this thesis. It has been observed that imitation seems to be by far a much more active area of research rather than other subtler forms of social learning such as Emulation, Stimulus or Local Enhancement and so on.

2.2.1 Programming by Demonstration

The concept of *Programming by Demonstration*, *Learning from Demonstration* or imitation learning in robots has been propounded early on. In fact, good surveys of work in imitative task learning were given by Billard in [8], Hussein [44] and Ravichandar [61].

Learning asocially, i.e. without any previous intimation or knowledge of problems that require attention to, is computationally expensive and energy inefficient. Programming by demonstration reduces search spaces by learning from examples that are already efficient. The robot can even eliminate bad search spaces by watching a certain solution fail. Further, it was opined that imitation is a natural way of learning from lay people, who are unaware of the intricacies of robot design. Lastly, imitation learning seeks to understand how perception can create successful actions in the perception-action coupling[8].

Historically speaking, programming by demonstration involved state-action-state sequences using primitive if-then-else rules[8]. To move forward from simply copying demonstrated movements to generalizing across sets of demonstrations, Machine Learning techniques were used. These techniques incorporated perceptual problems, such as the incorporation of vision stream replacing the term Programming by Demonstration with Imitation learning. [61] categorize demonstrations on the basis of how robots get data. They can get them through *Kinesthetic* teaching, i.e. users physically moving the robot, *tele-operation* and simple *passive observation*. Before imitation can be performed successfully, the following questions have to be answered: *What* to imitate, *who* to imitate, *when* to imitate and *how* to imitate ([8], [65]).

Apart from real instances of images/videos and speech that can be utilized by Machine Learning algorithms, Virtual and Augmented Reality can provide further data to learn from. In engineering approaches, a metric of imitation performance is defined, with different weights given to different skills, and an optimized controller seeks to reproduce the behavior the agent has observed. Furthermore, parameter search for encoding movement in joint space (16 dimensions for 16 joints, for example) is a very essential part of finding optimal controllers that can represent movements. Ideally, Imitation should help reduce this search space for initial states that lead to global optima fairly quickly[8].

Skill encoding can be done through various techniques. Statistical learning (which evolved into certain types of Deep learning), Hidden Markov Models (HMMs), and Dynamical systems such as Recurrent Neural Networks are but a few of these techniques[8]. Dynamical System Learning seems to be interesting for several reasons. The learning algorithm needs to be fast, and dynamical systems using techniques like locally weighed regression can perform one-shot learning. More work on one-shot learning using complex Deep learning algorithms has also been accomplished in [23] and [77]. Using attractor dynamics of nonlinear systems, it would seem that trajectories not just exactly the same as the demonstration, but also similar ones can be reproduced. Yamashita confirms this to a limited extent in [76]. This helps with the concept of generalization of movements and trajectory. It occurs that since similar movements can be categorized together, movement classification can also be done this way. In fact Principal Component Analyses for movement tracking using proprioception data (internal representations of the artificial neural network) in [76] and [45] confirms that this is possible.

Scaffolding and adding to previously learned behaviors is a concept that was proposed in both [65] and [8]. Incremental learning, as it is called, has the primary aim of making learning of demonstrations possible with fewer examples required. The concept of reusing primitives (lower level representations of movement) is stated as a goal several times previously, with Arbib proposing Motor Primitives for his Schema theory on a biological level[3][4], and [76] trying to reproduce it in Deep learning based Cognitive model for Robot control.

Such 'transfer learning' abilities have been demonstrated with Convolutional Neural Networks for image and video classification, text related applications using attention based networks for text classification and sentiment analysis, along with applications in biomedical engineering and bio-informatics [80].

2.2.2 Stimulus Enhancement and Emulation

Among the literature surveyed, very few actually allude to forms of social learning in robots other than imitation.

One of the few instances to the contrary was a study conducted by Cakmak et al. [14], where social learning forms other than imitation were compared. The study involved robot-robot interaction, where the social robot (learner) could implement one of four social learning mechanisms: Emulation, Stimulus Enhancement, Mimicking and Imitation, while learning from the demonstrator robot (social partner). The idea behind implementing such social learning mechanisms (other than imitation) is because they are simpler, computationally cheaper, and easier to implement.

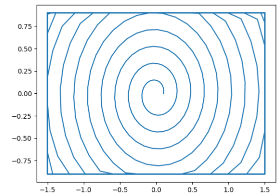
The goal of the study was to have robots learn how to make sounds with the help of one of two actions, grasping (with two hands) and shaking, or poking for different objects with different sizes.
For the demonstrator robot, learning was non-social in nature, and one of the robots observed the environment, approached the most salient object, performed a selected action, observed the effect, and went back to its initial position, updating the saliency of objects or actions (depending on the type of social learning)[14].
Learning from previous experiments[68], it was hypothesized that social learning would be helpful in cases where few objects make sound, i.e. when success was of lower probability.

2.3 Swarm Intelligence

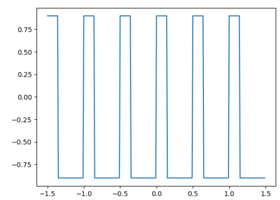
The idea to utilize several robotic agents that attract other agents based on certain information has been performed under a variety of different names. These agents that explore a static environment constitute algorithms that can be termed asocial or social in nature. Some previous work is mentioned below.

2.3.1 Asocial Exploration among Multi-agent Robots

It is important to understand the algorithms that have been used thus far to explore environments with some success. Gabal [30] and Fricke [27] utilize spiral patterns to explore an environment, with Fricke et al. extending the work to compare these patterns with ant foraging strategies, specifically the spiral based deterministic search, or the DDSA (Distributed Deterministic Spiral Algorithm). [20] utilizes another pattern wherein robots



(a) The archimedean spiral pattern used for exploration by all three robots



(b) Boustrophedon, or sawtooth patterns used for exploration

Figure 2.4: Patterns of search utilized for the Experiments

use a boustrophedon-like shape to search an environment. While this strategy covers more area, it also takes more time.

Search algorithms that have several agents search a space have been studied in detail. The observation that spiral search patterns for single agents are the most efficient has been extended to multiple agents using something called the Distributed Deterministic Search Algorithm (DDSA) in [27]. A simplistic scenario is to have several agents follow a spiral pattern at different paths in parallel. There is no communication between agents, and the agents follow a deterministic path. The study compared DDSA with a search pattern that was inspired by behavior found in ants (called Central Place Foraging Algorithm CPFA) and included local communication between agents. While the study was a simulation, DDSA was found to perform just as well or better than CPFA.
In a similar fashion, the idea of designed experiments is to compare a deterministic search pattern with one that includes communication and try to understand if a better methodology can be come up with that does not rely on a brute force search methodology.

2.3.2 Co-operative Foraging

Animal society based foraging techniques have been the focus of several studies over the past decades in multiple agent robotics. Pheromone based strategies include a wide range of techniques from actual chemical trails where robots rely on chemical sensors [29] to sense where earlier robots moved through, to vision based markers as pheromones [34] and virtual pheromones that only exist in computer memory stored and accessible through a central server [42]. A relevant honey-bee inspired algorithm (BEECLUST), which is inspired by the way honeybees navigate in a swarm, wherein they break into individuals under lower temperatures and congregate in higher temperatures along with some other rules was utilized in [66].

Strategies where one robot passes off relevant information have also been experimented with. Pitonakova [60] utilized what they termed 'recruitment' where a working robot, that spots an idle robot, recruits the idle one by sharing the location of the foraging site through direct communication. The adaptive behavior of the swarm is termed 'plasticity' and certain interesting observations are made. Recruitment of other idle agents is only useful for the swarm as a whole if the environment contains fewer resource caches that are harder to find, but once found have a large amount of foraging material in them. If the forage material is not of high quality, both recruitment and search for other resources becomes important. These conclusions have direct relevance to the study described in Chapter 3 and will be discussed further there. One key takeaway is that social learning might be just a 'mode' of operation, which may be switched on and off as and how it is considered useful for the swarm.

Some previous work on whether agents already in an environment can pass on the skills they have learned to a newer 'generation' of robots utilizing social learning was conducted by Noble and Franks [57]. Behavior Contagion and Emulation were used as social learning mechanisms in a simulation study where an agent could either follow a more experienced one for 25% of their life time, or a behavior contagion among the population with 10% probability the observer would do the same as the observer, or they could emulate the observer, i.e. they would know if there was payoff for a certain move, but not what the move was itself. Lastly, imitation meant the observer would copy both the move and the resulting reward of the demonstrator. Agents had a life time of 400 time steps and new ones would be born when older ones died. The environment had a very high reward density. Random moves were used to check if social learning mechanisms fared better than them. Mean payoff for social learning mechanisms was highest for the emulation model, and for observer's following the demonstrator for the first 25% of lifetime, mean payoff also increased. This situation can be likened to critical learning where social learning happens first. The experiment further shows that emulation or imitation of successful models provides the best payoff.

2.4 Connecting the Literature Review

After a brief introduction about the what and why of social learning, Section 2.1.3 surveyed the types of social transmission mechanisms that are used in both studies. Specifically, Local Enhancement is used for Study 1. The learning framework for this study in the context of Local Enhancement is identification of reward cache locations in the environment so that they can be used to forage from later on for self-sustainability. This does not consist of a belief-system such as a Bayesian learning which updates after reward locations are found. Concepts of Strategic Social Learning, i.e. when to learn are used for Study 1. Sections 2.3.1 and 2.3.2 then provided further background for Study 1 in the form of previous work done in multi-agent and cooperative robotics based foraging.

Observational Conditioning, Stimulus Enhancement and Response Facilitation are analyzed in Study 2. The background for Study 2 was covered in Section 2.1.3. Here, the targets of learning are the human participants, and both objective and subjective measures are used to attempt quantification of learning in human participants.

Chapter 3

Experiment 1: Social Learning in Multi-agent Robotics

> [Francis] Bacon in his instruction tells us that the scientific student ought not to be as the ant, who gathers merely, nor as the spider who spins from her own bowels, but rather as the bee who both gathers and produces
>
> Michael Faraday, The Life and Letters of Faraday, Vol. 2, p. 404, 1870

3.1 Research Questions

In order to establish Local Enhancement among a group of robots as outlined in Figure 1.1, the research questions for this study seek to understand if social learning can be helpful with time constraints, if reward density in the environment makes any difference to foraging socially or asocially, and how critical or conditional social learning might make a difference. Fundamentally, it would be good to understand if copying benefits all robots (as a society so to speak) or if certain robots are better placed to benefit the most. Through these set of experiments, we also want to understand under what conditions local enhancement would be beneficial to a robot society.

The research questions (RQs) are defined as follows:

1. How does the performance of social exploration in robots compare to asocial exploration under time constraints?

2. How does the performance of group exploration in social and asocial exploration scenarios compare under conditions of varying amount of rewards?

3. What are the effects of learning socially first (critical learning) as compared to learning socially later (conditional learning)?

Hypothesis 1 (H1) : Based on social learning theory, a simple social learning mechanism, when infused in a multi-agent robot society, should perform better than simple asocial explorers.

Hypothesis 2 (H2) : Critical social learning should perform better than conditional social learning in a stable environment (where reward locations do not change often) when primary reward density is low. This follows from discussion in Chapter 2.3.2 from earlier studies conducted by [60] and [57].

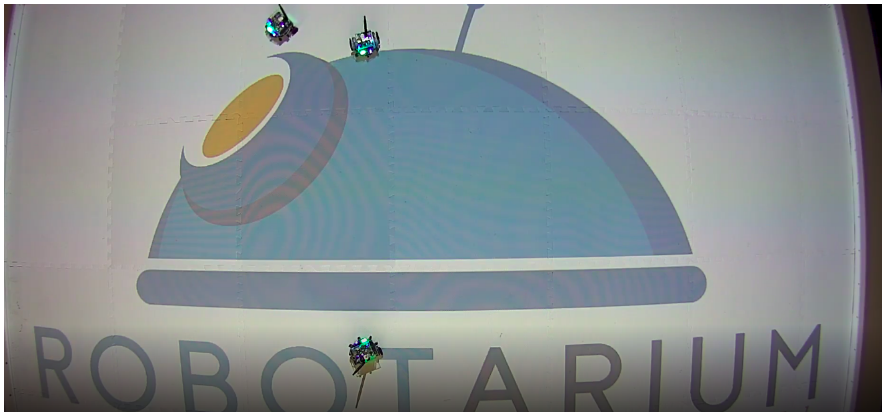
The focus on local enhancement follows from the discussion regarding computational complexity that involves an attempt at imitation. The focus for this study is on minimal real time computation for simple tasks such as foraging for survival. Further, to forage efficiently over larger areas, calling attention to the *location* of reward caches would be well suited to robots with limited computational capacity that are able to search a wider area with more numbers.

3.2 Methodology and Experimental Procedure

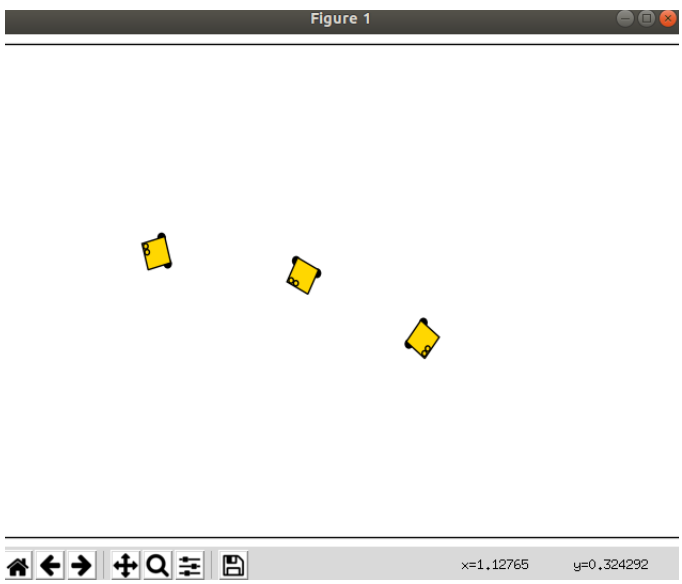
Because of the nature of the experiment, and because we would like for the robots to be minimalistic in both size and computational capacity, smaller swarm-like robots were utilized. While initially the Quanser based quadcopters were supposed to be used, due to COVID 19 shutting down labs, a remote platform called the *Robotarium* [59] based at the Georgia Institute of Technology at Atlanta, GA was used instead. The Robotarium allows jobs to be submitted via its portal online and accepts python based code[1]. The code first needs to be verified through simulation locally to make sure no explicit collisions take place, although safety certificates prevent actual collisions in the Robotarium arena. The simulation also provides for a rough estimate of how much time the expeirment is going to take. A representation of the simulation environment in Python on Linux is shown in Figure 3.1. Webots [52], a simulation based environment was also considered, but priority was given to real robots.

A group of three robots were used for all experiments with the understanding that in nature, packs of three foraging animals are a balanced number to hunt or forage [6]. The robotarium utilizes small robots called GITSBots which are 130*90*180mm in dimension in an arena of 3 * 1.8 m which represents the total search space. A representation can be seen in Figure 3.1. The robots contain ESP8266 micro-controllers that operate at 160 MHz. Rewards were simulated in computer memory and their locations recorded such that if the robot was within an *observational radius* 0.03 m of the reward, the robot would detect and capture the reward.
Reward Density refers to the total number of rewards present in the environment. These were arbitrarily selected to be $R \subset [5, 10, 20, 30, 40, 50]$ where R is the reward density. The environment here is considered *stable*, i.e. the reward cache locations do not change and their value does not decrease. It is to be noted that the Robotarium's clock speed is 0.033s, i.e. each 'iteration' in Robotarium time lasts 0.033s.

[1] The code is available online and more details can be found in Appendix C



(a) A still image from an overhead camera of the Robotarium running an experiment with three GritsBots



(b) The simulation environment

Figure 3.1: Patterns of search utilized for the Experiments

3.3 Experimental Design

For this exploratory study, to systematically address the research questions, a simple flowchart was designed to make the objectives of the various experiments clearer, and to show why they are needed. In figure 3.2, the left branch which starts with **Asocial Exploration** describes the three types of exploration algorithms that the first experiment explores while the right branch that starts with **Social Exploration** shows the different conditions that a group of robots can face while attempting cooperation for Experiments 2.1 and 2.2 .

The first experiment is designed to collect baseline statistics to understand which of these three approaches works best or in the most optimal manner when the three robots forage for reward caches. These three types of algorithms have been discussed in Chapter 2 and are regarded as more optimal than some of the other algorithms that were explored.

The workflow for experiment 1 for asocial exploration is established as follows:

- Asocial exploration is deterministic and involves robots exploring the search space in a deterministic fashion with preset patterns.
- Based on previous research [27], the parallel deterministic spiral formation is adapted as one search pattern. The sawtooth formation, also called the boustrophedon [20] search pattern, and randomized robot behavior (also tested in [20]) are the other two patterns compared for the baseline testing. These formations are visualized in Figure 2.4.
- The search space has different numbers of rewards, and constitute a total of six scenarios for each search pattern based on the six reward densities.

For experiment 2 involving social exploration, the workflow is as follows:

- Program a proximity check for *observational radius* to introduce social learning
  - A proximity check refers to checking surroundings to locate other robots that might be in a certain radius. In real life, these represent sensor limits
  - A broadcast mechanism that simply communicates success in finding the reward to any other agent in the proximity
  - The location of a primary reward is stored in an agent's memory such that it can revisit the site when it needs more food (rewards) to survive

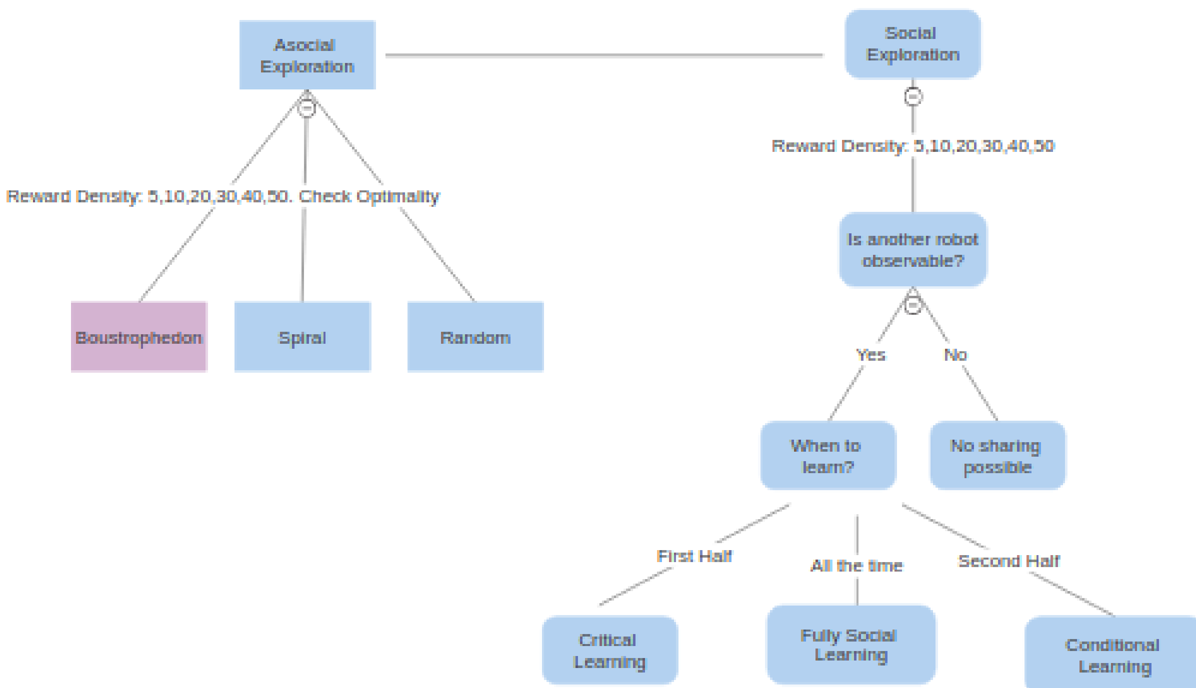


Figure 3.2: Flowchart outlining the three experiments for Study 1

- Implementing local enhancement
  - Local Enhancement is the redirection of attention towards a potentially interesting location that may or may not lead to social learning
  - For local enhancement to happen, the engineered solution that we use is to create a matrix of proposed locations of interest. These locations are shared socially using the above mentioned broadcast mechanism.

Because of the time based constraints, the following defines a robot's lifetime:

- For asocial exploration, an agent is provided with 6000 iterations (an iteration lasts 0.033s in real time in the Robotarium). Any reward location can only be awarded once to the agent. This is to maximize the understanding of how many such caches an agent can detect and forage for itself.
- An agent may explore socially the entire time. This mode is called *fully social* and the agent spends the entire time foraging and scrounging.
- For the *critical social* exploration experimental runs, an agent is given 6000 iterations, the first half of it spent exploring socially. The second half is spent exploring asocially.
- For *conditional social learning*, the agent spends the first half of its lifetime, i.e. the first 3000 iterations exploring asocially, followed by social learning in the second half. All other conditions are the same as critical social learning.

3.4 Experimental Procedure

3.4.1 Asocial Exploration

For experiment 1 as detailed in Table 3.1, we test the three conditions based on different search patterns to determine the most effective one under a time constraint. It is to be noted that Figure 2.4 represents the waypoints that each robots received parallel to each other, while the random condition simply gives the robots random waypoints one after another. All this is done with the environment having reward densities of [5,10,20,30,40,50], and the experiments are repeated 10 times for each condition, thus giving us 3 (robots) * 10 (repetitions) = 30 data points per condition. With 6 conditions in total, 180 data points were collected. The life-time of a robot was limited to 6000 iterations per robot. It is to be noted that these experiments were all performed on the Robotarium and took an estimated 240s (4 mins) on average per run.

Table 3.1: Types of Experiments

	Type of Experiment	Conditions	Aim
1	Asocial Exploration	Boustrophedon Spiral Random	Compare which asocial strategy forages/discovers most reward caches
2.1	Social Exploration	Fully Social Critical Conditional	Understand how sharing of rewards improves reward cache location by individual agents
2.2		Local Enhancement	Understand if an agent can successfully return to previously identified foraging site in 400 iterations

This table shows the experiments that were designed, what conditions were present for the experiments, and what their aims are.

3.4.2 Social Exploration

Two experiments were designed using social learning. The first (Expt 2.1 in table 3.1) dealt with the period of time agents should learn about the locations of reward caches from each other socially as opposed to not learning. The second (Expt 2.2 in table 3.1) was to answer the question of whether agents that already know of the presence of a reward cache either through socially learned or self-discovered reward caches can come back to these caches within a set time limit (400 iterations).

The same basic structure as was used for the asocial exploration was utilized for Experiment 2.1 with a few changes. Fully social learning was performed with all robots in the spiral formation (the reason for using spiral formation is explained in the Analysis section) and social reward-sharing happened when one robot 'observed' another pick a reward within an *observational radius* of 0.6 m as they were discovered.
For critical social learning, social learning with the same observational radius was allowed for the first 3000 iterations and asocial exploration for the rest of 3000 iterations. The opposite was done for the conditional learning condition, with the first 3000 iterations being asocial exploration.

Experiment 2.2 had a 'follow up' 400 iterations (13.2s) of life-time extension over the previous 6000 iterations if an agent found a reward cache earlier either socially or asocially. The objective was for the agent to find the most convenient reward cache to see if it could reach it within given time constraints so that they could continue foraging for a limited time with added life-time.

Convenience was defined as the closest (by euclidean distance) reward cache to the location of the agent at 6000^{th} iteration. The agent had at least one, but often multiple, reward cache locations in memory and would choose one from these *proposals*. A representative figure showing the proposals and the eventual closest solution that the three agents chose is shown in Figure 3.3

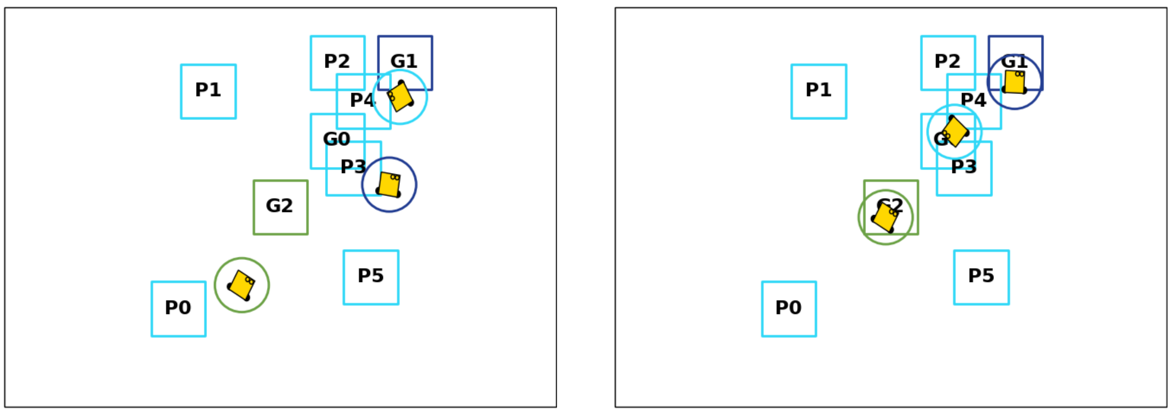


Figure 3.3: A demonstration for Experiment 2.2. The figure on the left shows Proposals (named P0-P4, in light blue squares) that are available to Robot 1 (with the light blue circle around it). The Robot heads to G0 (its goal) because it has found the locally enhanced region to be the most 'interesting' (since it is closest).
The figure on the right shows the agent successfully reaching its goal, and hence successfully realizing local enhancement by investigating a proposal that it found 'interesting'. Note that the other proposals were not investigated since they were not found 'interesting'.

The same set of reward densities were used for all the social learning conditions.

3.5 Results

3.5.1 Experiment 1

As stated in Section 3.4, the three patterns of asocial exploration were repeated 10 times each for statistical analysis to understand the differences between conditions in terms of means, variances and significance. These are described in Figure 3.4.

Mean	Social	Critical	Cond	Spiral	Random	Bous
5	2.067	0.733	0.933	1.000	0.200	0.200
10	2.667	3.000	1.733	2.200	1.267	0.867
20	6.867	5.000	4.733	4.267	4.600	1.600
30	9.467	5.067	7.800	6.533	5.400	3.133
40	14.733	10.133	9.867	8.000	6.867	2.800
50	15.600	9.067	10.600	8.800	9.867	4.533

Table 3.2: Mean reward cache locations foraged for all conditions in asocial exploration and social learning. The rows are formatted by strategies explained in Table 3.1 and the columns represent the reward density of the environment.

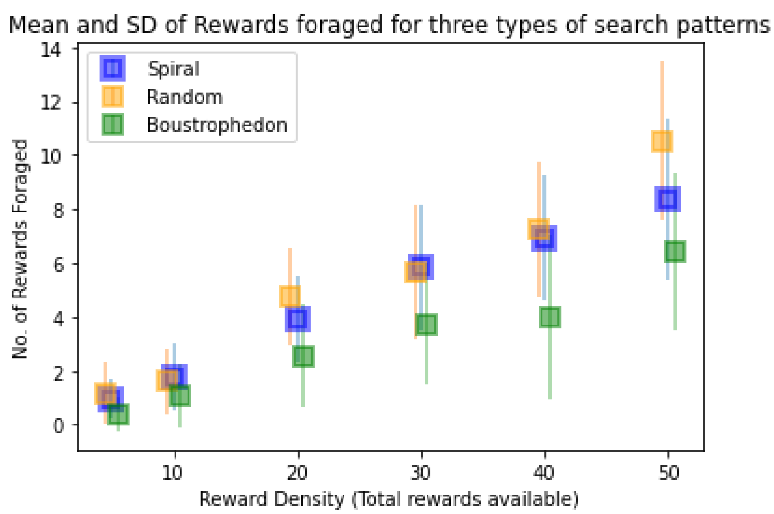


Figure 3.4: Means and standard deviations for all robots aggregated together to get a single value (Mean± Std Dev), for different asocial search patterns at each reward density for Experiment 1. Data points were shifted slightly along the horizontal axis to make the standard deviation bars clearer to read.

3.5.2 Experiment 2

As summarized in Table 3.1, Experiment 2.1 utilized different levels of social exploration across different environmental reward densities and yielded results which are summarized as means and variances in Figure 3.5. A comparison is also made with the Spiral asocial mode of exploration.

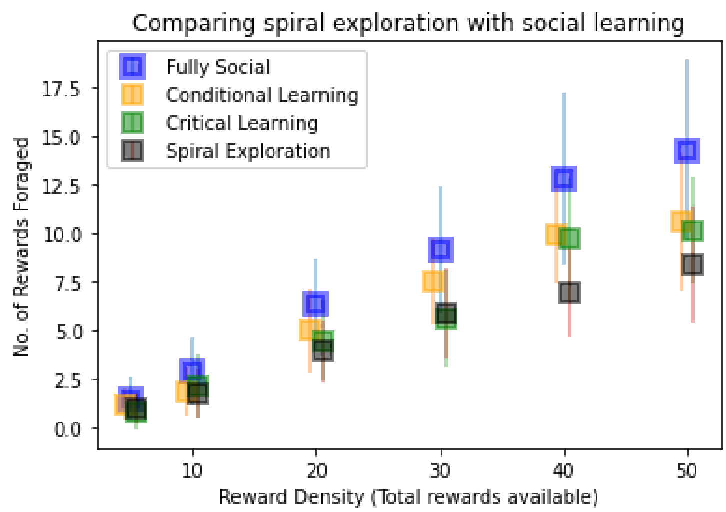


Figure 3.5: Comparison of means and standard deviations for reward foraging from three different types of social learning for Experiment 2.1. One selected type of asocial exploration (spiral) is also present for comparison. Data points have been shifted slightly to make the standard deviation bars clearly visible.

For experiment 2.2, Figure 3.6 represents the number of robots (3 robots per run * 10 repetitions = 30 per condition) that were able to locate and reach these reward caches (blue part of the bar) and those that weren't able to (red part) within the 400 iterations given to them. The red part also contains those agents that did not find a reward cache in the first place.

3.5.3 Data Analysis

Both experiments 1 and 2.1 were repeated such that they could be subject to statistical analysis to try and find if there were any significant differences within the conditions described. Experiment 1 had three conditions based on the type of asocial exploration

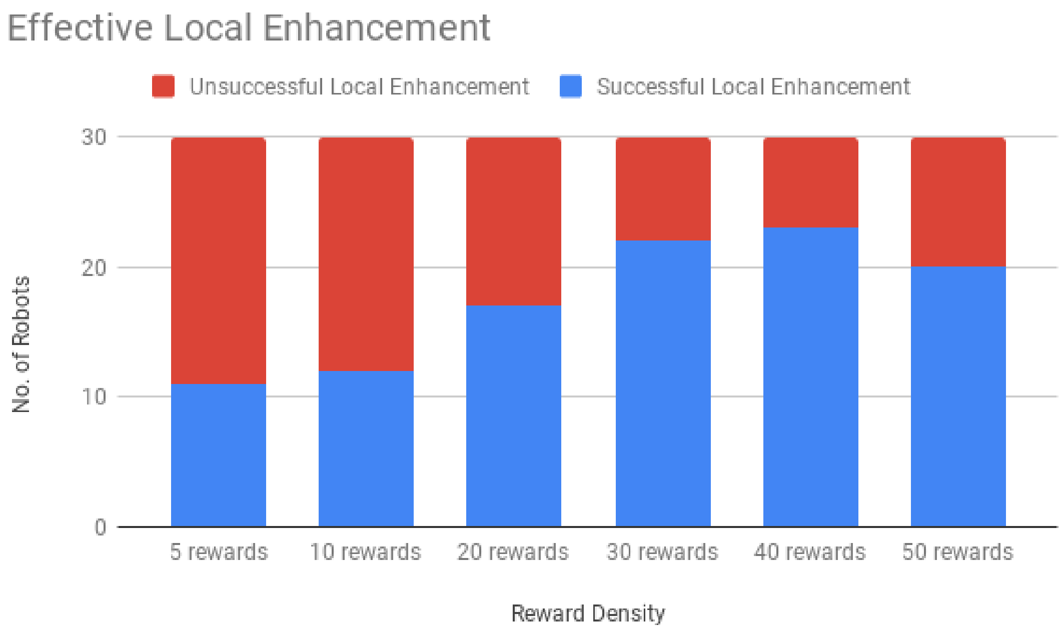


Figure 3.6: The bar chart shows how many robots across the varying reward densities were able to successfully locate the closest reward cache location and visit it so as to gather food/energy to forage for the next 400 iterations. This represents the results of the second social exploration experiment (Experiment 2.2) where robots are supposed to return to a reward cache that they have in their memory once they have exceeded the initial 6000 iteration steps (lifetime)

Table 3.3: Two way ANOVA for Asocial Exploration

	sum_sq	df	F	PR(F)
C(strategy)	444.4481481	2	48.47849829	4.88E-20
C(rew_density)	3698.209259	5	161.3539235	1.87E-103
C(strategy):C(rew_density)	176.5740741	10	3.851988577	4.73E-05
Residual	2392.833333	522		

Table 3.4: Post-hoc Analysis for Asocial Exploration

group1	group2	meandiff	p-adj	lower	upper	reject
Bous/20rew	random/20rew	2.1667	0.0121	0.2282	4.1052	TRUE
Bous/30rew	random/30rew	3.2667	0.001	1.3282	5.2052	TRUE
Bous/30rew	spiral/30rew	2.9333	0.001	0.9948	4.8718	TRUE
Bous/50rew	random/50rew	4.1333	0.001	2.1948	6.0718	TRUE
Bous/50rew	spiral/50rew	1.9667	0.0426	0.0282	3.9052	TRUE
random/50rew	spiral/50rew	-2.1667	0.0121	-4.1052	-0.2282	TRUE

algorithm. Experiment 2.1 had three conditions again based on the period of social learning. For both these experiments, two independent variables can be found. For experiment 1, these are the reward density and the type of asocial exploration. For experiment 2.1, these are the type of social learning and the reward density. In both cases, the dependent variable is the number of reward caches located. Thus, the two-way Analysis of Variance (ANOVA) was used to account for the two independent variables, with the Tukey's pairwise HSD as a post-hoc test to find which conditions specifically were significantly different from others.

Tables 3.3 and 3.4 describe the Two Way ANOVA and post-hoc pair wise Tukey's HSD for asocial exploration respectively.
Tables 3.5 and 3.6 then describe the Two Way ANOVA and post-hoc Tukey's test respectively for socially enabled exploration.

The Two-way ANOVA test results for asocial exploration with Reward Density and Asocial Exploration strategy (termed rew_density and strategy respectively) as the independent variables are shown in Table 3.3. The test results demonstrate that rewards that are foraged depend significantly on both the strategy and reward density, and that there is an interaction effect with $P < 0.05$

Table 3.5: Two way ANOVA for Social Learning

	sum_sq	df	F	PR(F)
C(strategy)	979.4819444	3	51.18205889	6.70E-30
C(rew_density)	9445.090278	5	296.1274597	1.32E-169
C(strategy):C(rew_density)	453.8597222	15	4.743216587	8.71E-09
Residual	4439.833333	696		

A follow-up post-hoc Pair wise Tukey's HSD test for asocial exploration with pairs of all possible combinations being compared against each other is shown in Table 3.4. Since there were over 150 pairs, only the most relevant data that differs significantly is shown in this table, i.e. those that reject the Null Hypothesis (hence reject = TRUE). The results confirm the hypothesis that Boustrophedon performs worst of the three strategies at reward densities 20, 30, 40 and 50, but while we can observe the means of random and spiral being higher at 5 and 10 reward densities, these differences are not statistically significant (i.e. $P > 0.05$). Interestingly, for the time limit given to the 3 robots, the random strategy seems to yield significantly higher reward locations than the spiral algorithm for the highest reward density of 50.

The Two way ANOVA test results with Reward Density and Social Exploration strategy (termed rew_density and strategy respectively) as the independent variables can be seen in Table 3.5. The test results demonstrate that rewards that are foraged depend significantly on both the strategy and reward density, and that there is an interaction effect with $P < 0.05$

A post-hoc Pair wise Tukey's HSD test confirms the hypothesis that fully social strategy performs the best of the three social strategies (significantly better than spiral, which is asocial but used for comparison, at 20,30,40 and 50, conditional learning at 40 and 50, and critical learning at 30, 40 and 50). There is no conclusive evidence that fully social performs significantly better at lower reward densities such as 5 and 10.

Table 3.6: Post-hoc Analysis for Social Learning

group1	group2	meandiff	p-adj	lower	upper	reject
conditional/40rew	social/40rew	2.9	0.0024	0.5183	5.2817	TRUE
conditional/40rew	spiral/40rew	-2.9333	0.0019	-5.3151	-0.5516	TRUE
conditional/50rew	social/50rew	3.6667	0.001	1.2849	6.0484	TRUE
critical/30rew	social/30rew	3.5667	0.001	1.1849	5.9484	TRUE
critical/40rew	social/40rew	3.1	0.001	0.7183	5.4817	TRUE
critical/40rew	spiral/40rew	-2.7333	0.007	-5.1151	-0.3516	TRUE
critical/50rew	social/50rew	4.1	0.001	1.7183	6.4817	TRUE
social/20rew	spiral/20rew	-2.4	0.0457	-4.7817	-0.0183	TRUE
social/30rew	spiral/30rew	-3.3	0.001	-5.6817	-0.9183	TRUE
social/40rew	spiral/40rew	-5.8333	0.001	-8.2151	-3.4516	TRUE
social/50rew	spiral/50rew	-5.8667	0.001	-8.2484	-3.4849	TRUE

3.6 Discussion

Research Questions 1 and 2 can be answered by comparing reward caches located as a group socially and asocially using Table 3.6. Fully social learning performs quite consistently better than any type of asocial exploration except at low reward density as can be seen from the significantly different results between social and spiral.

It is interesting to note that during asocial exploration at higher reward densities, for 20, 30 and 50 rewards in the environment, spiral and random patterns outperform Boustrophedon significantly. At higher reward densities, however, the random pattern outperforms both spiral and boustrophedon. This is unexpected and probably happens because of the limited time frame, since the boustrophedon is very thorough and covers the full search space if given enough time.

Lower reward densities are of special importance since this is where multi-agent robots should prove to be more helpful. Social learning does not seem to be *significantly* better than asocial exploration. However, the mean of reward caches located is higher quite consistently as shown in Table 3.2, and this can make a huge difference. Should these reward caches contain some of sort of resource that enables the robots to continue their search, social learning would mean that a group of robots could continue where the same group would be unable to if they did not share the resources they located.

Comparing table 3.2 with Figure 3.6, it is interesting to check if social learning gives a specific advantage to social learner agents for survivability by going back to reward cache locations that were already discovered. Because social learners at reward density of 5 and 10 locate double the number of caches compared to asocial explorers, they have a higher probability of coming back to a reward cache location, and thus increase their chances of continuing their work after their initial lifetime. This seems to confirm finding from [60], where a low reward density benefits from 'recruitment' as was defined there and recapped in Chapter 2.

As for Research Question 3, there are no significant differences between critical learning and conditional learning in the static environment. There seems to be little to no difference between the means, and this can be seen in Figure 3.5 and Table 3.2. Therefore, for Research Question 3, no significant differences can be found and this experiment does not provide any conclusive evidence that learning first or learning later makes any difference when it comes to a stable environment.

A good reason why switching between social and asocial exploration can be important is because in the scenario where the robotic agent is exploring an environment with other robots but does not know what the reward density of the environment is, it would make sense to share cache locations socially. However, sharing constantly would require battery life (transmitters and receivers have battery life costs associated to them) and so sharing continuously becomes more of a burden. In this scenario, an instinct switching system that decides when switching between social and asocial exploration is optimal could help make these groups of robots be more energy efficient and robust to challenges related to the environment in terms of finding resources.

3.7 Conclusions, Limitations and Future Work

An exploratory study was designed to understand if social learning provides benefits to a group of robots foraging for reward under different reward densities and modes of operation. Asocial exploration proved less effective than social exploration overall, and while differences between them were not significant for lower reward densities, social exploration would certainly be more helpful in keeping a group of robots 'alive' longer if they had to forage for their own food. The environment considered here was stable, i.e. the locations of caches did not change, and more work needs to be done on 'unstable' environments where the location and values of the caches change. Moreover, while switching between social and asocial exploration modes here happened deliberately at the halfway point, an algorithm that decides when to switch between the two modes on the basis of an estimate of the environmental reward density would help make this process more effective and optimized.

Extending the current work to more than 3 agents is also an important step with different reward densities. Lastly, due to the low availability of the Robotiarium during COVID19, it was not highly practical for more than 10 repetitions per case. More data might allow us to look at further trends.

This work was meant to lay the foundation for a research program into engineering social behavior among groups of physical robots working together. Providing robots with Computational intelligence, while making rapid progress, is still quite limited if robots are to be produced cheaply and en masse, which supports our research direction to investigate simple mechanisms of social learning that could run on relatively simple agents. For future work, the next step would be to use fully autonomously running robots, as soon as our robotics laboratories re-open.

There are further concepts from animal behavior that can be incorporated as social behavior in a group of robots. Positive and negative rewards can be used to induce location enhancement and avoidance. The current study focuses only on positive rewards. Avoidance can be induced by introducing history (previous experience) along with the option of sharing such information for negative rewards (punishment), which might create a culture of avoidance in the robot society. This would require ML techniques to take history into account, such as regression trees or Recurrent Neural Networks.

3.8 Transitioning to the Second Study

This study was meant to explore social learning between robots, or Robot-Robot Interaction. Difficulties with resuming experiments with physical robots (either aerial drones or small autonomous robots) due to COVID 19 meant we could not continue this experiment as the Robotarium was not readily available all the time. The Robotarium also had problems in that it did not allow any modifications to the environment provided (due to the remote nature) and no actual sensors existed on the robots.

Therefore it was decided to explore social learning in Human-Robot Interaction. Again, due to Covid 19 this study could not be carried out with physical robots and was conducted through an online study described in Chapter 4.

Chapter 4

Experiment 2: Social Transfer of Information using Virtual Agents

> The most exciting phrase to hear in science, the one that heralds new discoveries, is not 'Eureka!', but 'That's funny ...'
>
> Isaac Asimov, 1987.

4.1 Motivation

As explained in the previous chapter, this study focuses on Human-Robot Interaction with the robot utilizing simple forms of social transmission of information to understand if successful social learning can indeed happen between robots and humans. While the experiment was originally designed to constitute Human - Multi-agent interaction keeping in line with earlier research utilizing multiple agents, this study became more of scientific exploration since previous work on social transmission of information from robots to humans, virtual or real, has not been explored, specifically with the use of methods of transmission other than imitation. The study therefore establishes basic principles that can be used for further exploration in this domain.

This specific experiment is of great interest to us for a variety of reasons. New environments pose a challenging problem when humans who are naive to the environment first

Table 4.1: Description of similarities and differences of Study 2, as compared to Gerull and Rapee [35].

	Gerull et al.	Current
Mean Age	17 months	30.6 years
Type of Social Transmission	Observational Learning	Observational Conditioning
N - Participants	30	44
Gender of Observers	F-15, M-15	F-17, M-27
Demonstrator	Mother (Human)	Robotic Agent
Stimulus	Snake/Spider toy	Virtual Alien Animal
Measurement	Subjective -Likert Scale	Subjective and Objective
Number of times observer is exposed to stimulus	3 times 1st at 1 min, 2nd at 2 min, 3rd at 10 min	2 times 1st at 1 min, 2nd at 7 min

encounter challenges. An experienced agent that understands the subtleties and problems of an environment can pass on information to newcomers, gained by previously exploring this environment. We therefore envision situations when embodied artificial agents, quite specifically robots, introduce humans to a new environment efficiently. For instance, teaching humans about features and situations in the environment that are dangerous, and others that are harmless, even if they may appear dangerous.

In order to closely follow the scenario above, the experiment that comes closest involves a mother transmitting social information to a child about a novel stimulus in the environment as was described in Chapter 2.1.3 from Gerull and Rapee [35]. There are some major differences between what we proposed for this second study and what was done by Gerull and Rapee. First, grown adults do not necessarily react to surprising stimuli in a new environment in a manner similar to how toddlers or infants react. For example, adults may have inherent biases from previous experiences that might make them more or less susceptible to either positive or negative actions. Second, adults are already used to a large variety of different objects and experiences, so it can be difficult to create truly novel stimuli that they previously have never encountered, making our task of replicating the experiment difficult. Table 4.1 outlines the similarities and differences in our approach as compared to [35].

In this study we are interested in understanding if social information (in this case affective behaviour by a virtual robot) can spread to humans who share the environment with the virtual robot. It is likely that the social transmission mechanisms involved are types such as Observational Conditioning, Stimulus Enhancement and Response Facilitation, all of which were discussed in 2. Figure 2.3 portrays the decision making process that allows us to narrow down these mechanisms.

The study was planned to be completely online served as an immersive game environment to the participant who plays from a first person perspective. In order to design this study, a novel stimulus in the form of an alien animal that appears after a certain amount of time in order to startle the participant. A demonstrator in the form of a robotic agent introduces the participant to the world and shows the participant around. The robotic agent then encounters and reacts either positively or negatively to the animal, i.e. acting in either a friendly manner or running away in terror, respectively. All the while, the participant/observer watches the exchange without being able to move. The study used a between-subjects design, i.e. participants saw either the demonstrator's positive or negative reaction.

In this context, and with reference to the Decision Tree in Figure 2.3, the effect is stimulus specific, sensitive to the demonstrator's actions, action specific, but not a novel action sequence[1]. Thus, we are left with three possibilities to classify the social transmission that might occur in this study, namely, ORSL (Observational Response-Reinforcer Learning), RF (Response Facilitation) + OC (Observational Conditioning) or RF + SE (Stimulus Enhancement) which is the red box, second from left in Figure 2.3. In order to differentiate between these 4 types of social transmission mechanisms, two factors are important; whether the observer observes the demonstrator get rewarded for its behavior, and whether this learning is S-S (Stimulus- Stimulus) or R-S (Response-Reinforcer).

Stimulus Enhancement (SE), Observational Conditioning (OC) and Response Facilitation (RF) do not typically involve rewards[2]. Observational R-S Learning can therefore be ruled out. This narrows down the possible social transmission mechanisms to either RF + OC, RF + SE, just RF or SE. While this discussion is by no means complete, and more information is provided once the results are explained later on, it does provide us with a likely hypothesis about the types of social mechanisms at play.

[1]We assume that adult participants are fully capable of recognizing the demonstrator's actions. Hence, this effect is not based on a novel action sequence.

[2]There is some debate regarding usage of reward for OC, see [58]. However, OC combined with RF may not necessarily involve rewards

Figure 4.1: Alien Animal

4.2 Research Questions and Hypotheses

A between-subject study was designed that aims to investigate the effect of a *robotic* demonstrator on an observer[3] in an unknown terrain, i.e. a virtual forest. The behavior from the robot is shown as one of two affective states, positive or negative emotions expressed through body movements, portrayed by the robotic agent when encountering the novel creature, henceforth called the alien animal. Note, in order to provide context for the demonstrator agent's skills, it is introduced as having prior knowledge of the environment, which is new to the observer. Our research questions were as follows:

- RQ1: Does the perceived reaction (positive, negative) of the demonstrator (robot) affect the observer's (human participant's) response? If so, to what extent?
- RQ2: Does the participants' gaming experience or gender affect their responses towards the creature? If so, to what extent?

Reviewing the available literature, we propose the following hypotheses:

[3]We use the word observer and participant interchangeably since in this experiment, the observing agents are the human study participants.

- **H1 for RQ1**: Distance of the participant to the animal in the positive condition should be lower than in the negative condition. We expect participants in the positive condition to have a more positive attitude towards the alien animal. For example, in [35], children observe and infer the nature of the relationship between their carer and the novel creature by recognizing disgust or fear, and replicating the same avoidance behavior towards these creatures.

- **H2 for RQ1**: Participants in the positive condition should have a more positive and less violent perception of the creature as compared to the negative condition.

For H1, data was collected in the form of objective position which included measuring the position (in 3 dimensions in the game environment) of the participant and calculating the Euclidean distance to the animal once every second. Data collection for H2 was through subjective measures in the form of questionnaires of participants' perceptions of the alien animal.

4.3 Method

Since it was not possible to conduct in-person studies, due to COVID-19 based restrictions, a virtual experience was designed to emulate real world conditions as best as possible. The virtual experience was designed through Unity and served through WebGL to the participants. We used Amazon Mechanical Turk to carry out this remote study. The virtual environment is set in an unknown terrain (a forest). The game scene includes a virtual robotic agent capable of two types of emotional body expressions (positive, negative) portrayed by the robotic-like agent when subjected to a novel creature, with the robotic agent being depicted as someone with experience in the environment through the game narrative.

For the participants, the goal of the game is to collect spheres scattered around the game world. The spheres were distributed around the environment in such a way that it would not bias participants towards either moving towards or away from the alien animal. The entire participant experience was divided into three parts:

- The pre-game Questionnaires: This includes some basic demographic information, consent and general information regarding the game, the Ten Item Personality Questionnaire (TIPI) [39] and the Fear Schedule Survey (FSS) [5]. This was administered using Qualtrics© (2021 Qualtrics).

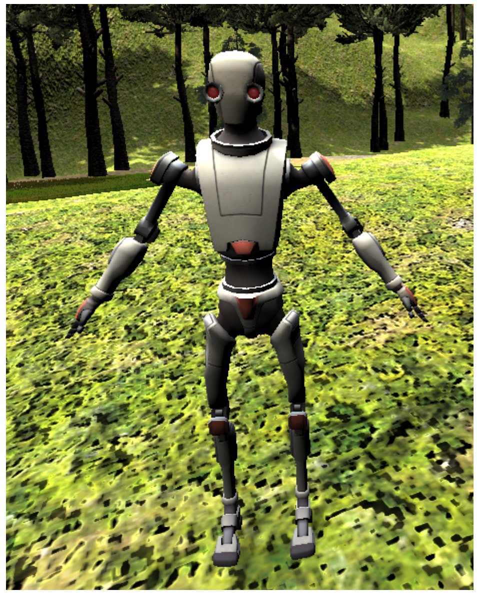


Figure 4.2: Virtual Robotic agent used.

- The game experience included three levels:
  - Level 1: Familiarization phase where participants get used to the game interface using their keyboard and mouse on a PC/Laptop
  - Level 2: Participants explore the area, looking for in-game rewards. The event with social transmission where the robot reacts either negatively or positively to the alien animal happens then. Next, the animal is removed from the game (in accordance with the methodology followed in [35]). This level then continues for another 6 minutes, to make sure there is no immediate emotional carry-over to the next phase.
  - Level 3: The participant starts the game at a position where they can observe the alien animal directly. The alien animal remains at the same position and does not move in space, it simply looks around. The participants can now explore the area in the vicinity of the animal, as well as moving further away, picking up resources. This level lasts 7 minutes before the game is bought to a close.
- The post-game Questionnaires: A set of custom-made questions ask about participants' experiences during the game, notably their perception of the alien animal, and

Figure 4.3: Flow of Experiment 2

including a few questions as attention-checks. This questionnaire is also hosted on Qualtrics.

The flow of the experiment can be seen in Figure 4.3. It is to be noted that the behavior of the alien animal stays the exact same in both scenarios. The only part that differs is the robot's reaction to the animal.

The pre-game and post-game questionnaires consist of questions that have been detailed in Appendix B. A preview of the game play is also given here.

This study was approved by the Office of Research Ethics and all the relevant parts can be found in Appendix A .

4.3.1 Game Development

The game was designed with the Unity 2019 editor. Participants were asked to enter their ID, and their in-game positions were recorded once every second. The game is hosted through the Web Graphics Library (WebGL) format on Github pages. This allows the game to be hosted remotely and to be playable on a web browser without the user having to install any plugins, and multiple participants can play simultaneously. In-game data is stored on an external online DataBase based on a MongoDB server on the cloud. The WebGL platform can be seen in Figure 4.4.

4.3.2 Questionnaires

Two standardized and two custom-made questionnaires were included in the study and administered through the University of Waterloo's Qualtrics system, along with the accompanying consent form at the beginning, an explanation of the tasks, instructions to play the game and a web page explaining the exact purpose of the game at the end of the experiment. Custom-made pre-game and post-game questionnaires were completed by participants. Attention check questions were used to address some of the issues regarding

Figure 4.4: The WebGL interface that participants used to play the game

Amazon Mechanical Turk studies such as lack of attention or misrepresentation [51] (e.g. Have you encountered the creature? How does the creature look like?). Other questions concerned participants' reaction to and perception towards the alien animal in relation to what they observed from the virtual robot demonstrator (e.g. How did you perceive the behaviour of the robot?).

4.3.3 Participants

A total of 49 successful participants were recruited on Amazon Mechanical Turk. Data integrity was important since there was some technical data loss through the three levels due to packet data loss for the participants' position which was to be transmitted to the MongoDB Atlas VM. In order to enhance the quality of the data collection, participants were recruited with respect to specific metrics ($\geq$ 96% completion rate for at least 50 previous tasks completed).

The participants were equally distributed to the two experimental conditions. However, due to some participants failing the attention checks, the study finally had 27 participants in the negative condition and 22 in the positive condition. Of the 27 in the negative condition, 3 were excluded due to packet data loss regarding their location in-game. A further 2 were excluded due to low quality of submissions. A low quality submission was defined as the participant either standing at the same place for a long time (over 3 minutes of the total 6 minutes in the first level, hence showing inactivity), or answering the questionnaire by clicking on the same options for multiple questions. This left us with 22 participants for each condition, in total 44 participants, 28 participants self-identifying as men and 16 as women, 19 to 55 years of age, with the mean age being 30.64, S.D. of 8.9 years, 32 gamers and 12 non-gamers.

4.4 Data

To understand the effect of social learning mechanisms on the participants, we collected the data in the form of objective and subjective measures.

4.4.1 Objective Data

The objective data collected during the study in the second phase includes mean and absolute distances [4] from the participant to the alien creature, the frequency of returns of the participant (moving back towards) to the creature, and the time duration that the participants took to their first return to the creature. A 'return' is defined as the participant walking to a position where they can clearly observe the animal from a distance of 40 game distance units. 40 is chosen as the cut-off since the landscape consists of hills near the position of the creature where a participant might be behind one of them and therefore not directly have an animal in their line of sight despite close proximity. These distances are calculated for each participant in relation to the creature at different time points: 15 sec; 30 sec; 45 sec; 60 sec and 100 sec. Table 4.2 describes the normality tests that were performed to determine whether a non-parametric or parametric test was required to test for significant differences between the positive and negative conditions. Shapiro-Wilk tests were performed to understand normality of the objective data. This test checks whether a distribution is *non-normal*, i.e. if $p<0.05$, the data is *not* normal. If the data is indeed non-normal, the Mann-Whitney U test was performed, and the independent-samples T-test on the data that were normally distributed.

4.4.2 Subjective Data

The subjective data includes a question regarding the participant's perception of the behaviour of the creature. The question that the participants were asked was: *How did you*

[4] For mean distance, we averaged the distance of a participant to the animal up to a certain time interval, for example, 30 seconds. The absolute distance is the distance to the animal at that point in time, for instance, at 30 seconds.

Table 4.2: Mean and Absolute distances of participants to the animal at 15,30,45,60 and 100s. For Absolute distance at 60s, where the positive was normal and the negative non-normal, the Mann-Whitney was used, since both distributions need to be normal for the T-test. Here N- refers to Non-Normal, and N+ to Normal

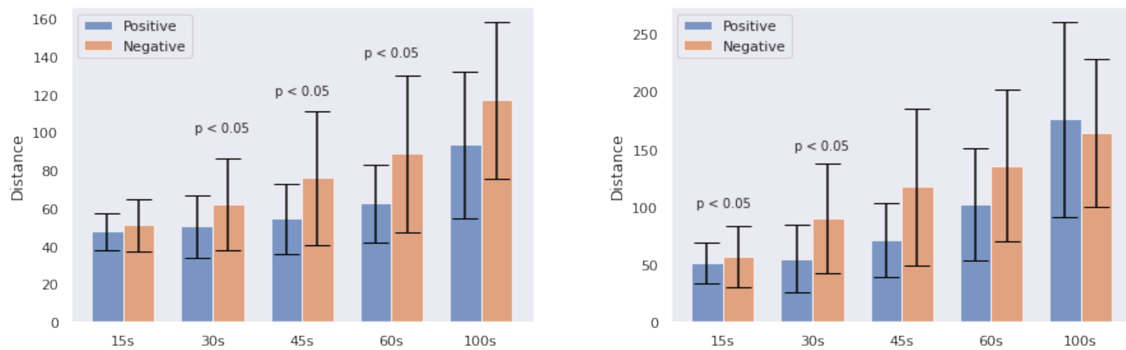
		15s	30s	45s	60s	100s
Mean	Positive	p<0.05, N-	p<0.05, N-	p>0.05, N+	p>0.05, N+	p>0.05, N+
	Negative	p<0.05, N-	p<0.05, N-	p>0.05, N+	p>0.05, N+	p>0.05, N+
Absolute	Positive	p<0.05, N-	p<0.05, N-	p>0.05, N+	p>0.05, N+	p>0.05, N+
	Negative	p<0.05, N-	p<0.05, N-	p>0.05, N+	p<0.05, N-	p>0.05, N+

perceive the behaviour of the creature?. Two example responses were: "I perceived the behaviour as curiosity" and "It seemed violent". The responses of all the participants were classified into three categories of perceived threat: Mild, High and None. Two independent coders not involved in the study rated the responses of the participants according to the three categories. Cohen's κ was run to determine if there was agreement between the rater's judgement on whether the participants perceived high, mild or no threat from the alien animal. There was good agreement between the two judgements, $\kappa = .726$ (95% CI, .556 to .896), $p < .001$.

One other pertinent question asked in the post-game questionnaire was about the participants' reactions when they *first* saw the creature. This is in reference to Level 2 (see Section 4.3) when the participants first saw the creature and the robot's reaction to it. The responses were one of [Run Away, Approach, Neither, Other] and this data was classified as Nominal, and therefore the Chi-Square test was used with z-tests for post-hoc analysis.

How did you perceive the behaviour of the creature?	I perceived the behaviour as curiosity
	It seemed violent
When you first saw the creature, what was your reaction?	I should run away
	I should approach it
	Neither
	Other

Table 4.3: Relevant perception based questions for subjective data



(a) (b)

Figure 4.5: a The mean distance of participants to the alien animal up until 30s, 45s and 60s, respectively, showed a significant difference between positive and negative conditions. b The absolute distance of participants to the animal up until 15s and 30s showed a significant difference between positive and negative conditions.

4.5 Results

4.5.1 Mean Distance

To calculate statistical significance for mean distances between participants and the alien animal, the Mann-Whitney U test was run to determine if there were differences in mean distance score between positive and negative conditions. Mean distance score at time point 30 sec, was statistically significantly higher in the negative condition (Mdn = 57.58) than in the positive condition (Mdn = 48.40), U = 152, z = -2.113, p = .035, using an exact sampling distribution for U (2-tailed). Similarly, at time points (for Mean distance) 45 sec and 60 sec, a Welch t-test was run to determine if there were differences in distance score between the positive and negative condition.

The mean distance for the participants from the creature in the negative condition was higher (at 45 sec: M = 76.07, SD = 36.11; at 60 sec: M = 88.91, SD = 42.43) than the participants in the positive condition (at 45 sec: M = 54.49, SD = 19.00; at 60 sec: M = 62.64, SD = 21.12), a statistically significant difference was found for both time points 45 sec, M = 21.58, 95% CI [-39.31, -3.85], t(31.79) = -2.481, p = .019, d= -.52 (Moderate effect) and 60 sec M = 26.27, 95% CI [-46.89, -5.65], t(30.80) = -2.6, p = .014, d= -.55 (Moderate effect).

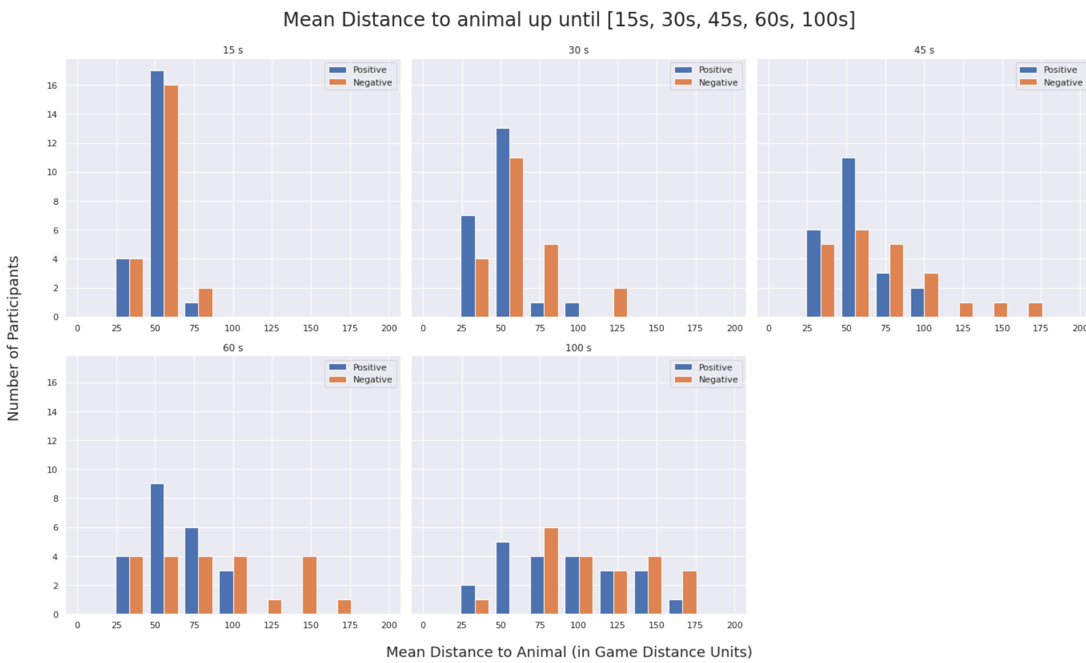


Figure 4.6: The histograms present the number of participants (y-axis) in a certain distance range (mean distance up to a certain time) to the alien animal (x-axis). In general, it can be seen that participants in the negative condition (orange) are more spread out and to the right (further away from the animal) whereas participants from the positive condition (blue) are not as spread out and favor the left side (closer distance to the animal).

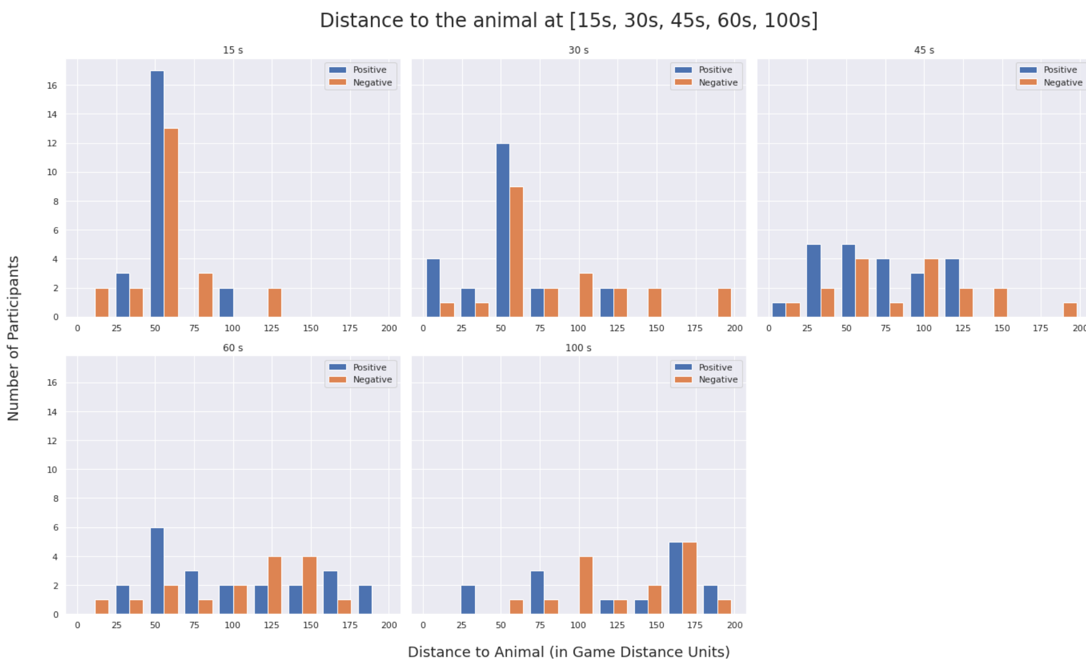


Figure 4.7: Exact distance (x-axis) participant/alien animal, and how many participants were in that range (y-axis). While participants in both conditions at 100s and 60s are quite spread out, from 45s and below, the participants in the positive condition (blue) are more heavily skewed towards the left progressively. This shows they are closer to the animal overall than their counterparts in the negative condition at that point in time.

4.5.2 Absolute Distance

To calculate if there are any significant differences in the absolute distance between participants and the animal, a Mann-Whitney U test was run to determine if there were differences in absolute distance score between positive and negative conditions. The absolute distance scores at time point 15 sec were statistically significantly higher in the negative condition (Mdn = 56.36) than in the positive condition (Mdn = 46.75), U = 153.50, z = -2.078, p = .037, using an exact sampling distribution for U (2-tailed). The absolute distance scores at time point 30 sec were statistically significantly higher in the negative condition (Mdn = 70.08) than in the positive condition (Mdn = 51.06), U = 175=4.00, z = -1.596, p = .013, using an exact sampling distribution for U (2-tailed).

4.5.3 Perceptions of participants towards the behavior of the animal

With respect to the first impression of participants towards the creature, a chi-square test of homogeneity was run, with an adequate sample size established according to Cochran[18]. The two multinomial probability distributions were equal in the population, $\chi^2(3) = 8.581$, p = .035. Participants in the negative condition were more likely to respond with 'Run Away' (n = 19, 86.4% versus n = 12, 54.5%). Post hoc analysis involved pairwise comparisons using multiple z-tests of two proportions with a Bonferroni correction. Statistical significance was accepted at $p < .0125$. With the Bonferroni Corrections, there were no statistically significant differences in any of the cases. While the 'Run Away' response for Positive vs Negative Condition is statistically significant *without* the Bonferroni correction ($\chi^2(1)$=.021), with the correction this is still greater than .0125.

A Mann-Whitney U test was run to determine if there were differences in subjective rating scores between positive and negative conditions for the question "How did you perceive the behavior of the creature?". No significant results were found for either of the raters' scores between Positive and Negative conditions.

4.5.4 Gaming Experience

Mann-Whitney U tests were run to determine if there were differences in the distances between Gamers in the positive and negative conditions. No statistically significant differences were found for either Mean or Absolute distances. This was extended to Mann

Whitney U-tests between gamers and non-gamers without regard for the condition they came from. No significant results were found for this test either.

A Mann-Whitney U test was also run to determine if there were differences in subjective ratings score between Gamers in the two conditions. No significant results were found for either of the raters' scores between the two conditions. Further, subjective rating scores between gamers and non-gamers were also non-significant.

4.5.5 Effect of Gender

The Mann-Whitney U tests were run to determine if there were differences in the distances according to gender. None of these data points for either Mean or Absolute distance were found to be statistically significantly.
Further, the Mann-Whitney U test to determine if there were differences in subjective rating scores between the self-reported Male and Female genders (no other genders were reported) found no significant results for either of the raters' scores between Male and Female participants.

4.5.6 Frequency of return to animal

An analysis was conducted to understand whether participants engaged in searching the area around the animal for rewards and returning to it. As before, the definition of 'return' was given as approach of participant to within 40 distance units, i.e. within observable distance.

A Mann-Whitney U test was run to determine if there were differences in the number of times participants return to the animal in the Positive and Negative Conditions. The statistic for frequency of returns was statistically significantly higher in the negative condition (Mdn = 1, Mean Rank=19.27) than in the positive condition (Mdn = 2, Mean Rank=25.73), U = 171, z = -1.722, p = .045, using an exact sampling distribution for U (1-tailed).

4.6 Analysis & Discussion

It is essential to interpret the results detailed above within the context of the definitions provided in Section 2.1.3 and understand if the results support our stated hypotheses within the definitions laid out earlier.

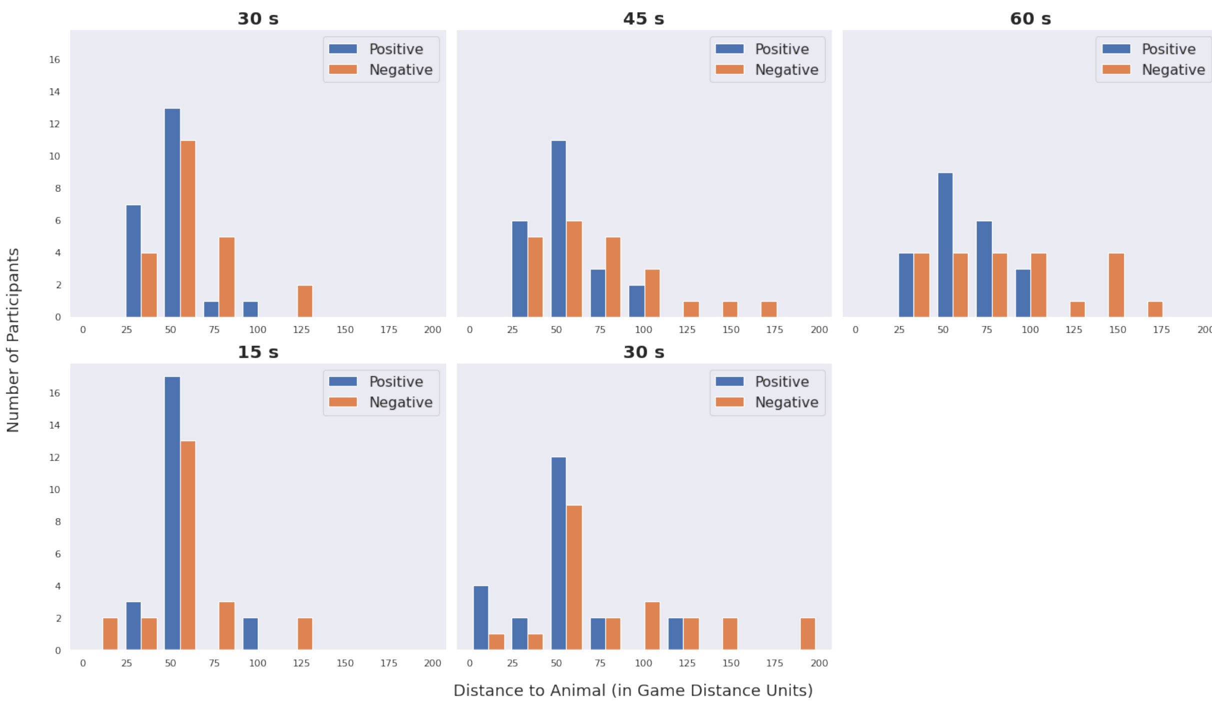


Figure 4.8: Row 1: The mean distance of participants to the alien animal up until 30s, 45s and 60s, respectively, showed a significant difference between positive and negative conditions. Row 2: The absolute distance of participants to the animal up until 15s and 30s showed a significant difference between positive and negative conditions.

4.6.1 Objective Measures for Positive & Negative conditions

A closer look at the histograms in Figure 4.8 provides an interesting trend where we observe the distance of participants in positive condition (blue bars) to be less spread out and closer to the animal (higher towards the left side of the histograms), which represents a lower distance to the alien animal. This is visible more clearly for 30s, 45s and 60s in Row 1. The differences in the mean distances can be seen more clearly in Figure 4.5a that shows significant differences between the two conditions.

A similar trend emerges when we compare absolute distances between the two conditions. Histogram bars in Row 2 of Figure 4.8 show that participants in the positive condition are in higher numbers at shorter distances from the animal, i.e. there are more participants closer to the animal in the positive condition for at least 15s and 30s. This trend disappears as time stretches further. This means they tend to be closer to the animal during the earlier stages than the participants in the negative condition. We hypothesize that this is because participants start exploring the area more randomly in search of rewards as time goes on. The trend of significant differences between participants in the two conditions is shown in Figure 4.5b.

One other point of interest is the result from the frequency of number of returns. This frequency was calculated to provide an understanding of how comfortable participants felt in searching for rewards closer to the alien animal. Assuming participants from the positive condition return to the animal more often due to having been exposed to a prior positive relationship, the 1-tailed test is significant. This *confirms Hypothesis 1 for Research Question 1*, i.e. the distance of the participant to the animal in general is significantly lower in the positive condition than the negative.

4.6.2 Subjective measures

The subjective data that was collected with respect to independent coders' ratings of the threat level that participants expressed towards the animal was non-significant. While there is a difference between the means of the ratings given by the coders for the two conditions, no significant results could be found. We hypothesize that this is due to hesitance among participants to admit fear in general, and they may not agree that they felt frightened even if their in-game behavior shows this. This means that *Hypothesis 2 for Research Question 1* cannot be verified. Overall, while the results suggest that there is an effect of the demonstrator's (robot) reaction on the participant, these results are not significant. We consider subjective data to be supporting data, and secondary to the objective data.

4.6.3 Gender and Gaming Experience

A close look at the relevant Sections regarding previous gaming experience and effect of Gender in Section 4.5 show us that neither gender nor prior gaming experience made a significant difference in participants' attitudes towards the alien animal. Since no differences can be found between gamers and non-gamers for either the subjective or objective data, our data suggests that gamers and non-gamers have a similar experience in our study. Furthermore, gamers in the positive and negative conditions also do not seem to have had a very different experience. This is suggested by neither the objective nor subjective data for the positive or negative conditions being significantly different.

4.6.4 General

Taking the three types of objective data measures together suggests that participants in the positive condition tend to stay closer to the animal, and the only difference between the games played by participants in the two conditions is the type of reaction to the demonstrator.

With this in mind, referring back to the literature in social transmission among humans and other animals, Heyes' definition of Stimulus Enhancement [41], where an observer observing a stimulus at time t1 causes an observable change in the observer's behavior at time t2, is satisfied. Further, it can be argued that the observer participants are exposed to a stimulus relationship (fear or positive reactions towards a stimulus, i.e. an alien animal) at t1, which then leads to a similar manner being adopted by the observers at t2, thus fitting the definition of Observational Conditioning as well. To provide support that Response Facilitation has happened, we have to show that the probability of an observer doing the exact same thing that it saw the demonstrator doing must increase. Two problems can arise while trying to conform to this definition. First, we cannot calculate the probability of adoption of the same action because we must first define what the 'same' action is. Second, we must demonstrate such a probability increases, which we cannot do since the repertoire of actions that the participant can take through the game is very limited. Hence, it is difficult to provide support for RF through our online experiment, and so this theoretical issue remains inconclusive and has to be investigated in future research.

Of particular importance to theories of fear acquisition, negative stimuli pairing created a negative impression that lasted over 6 minutes, which was the time given between Level 2 and Level 3. This means that the effect lasts at least 6 minutes and does not simply disappear after the initial appearance of the alien animal.

As a comparison to the earlier study by [35], this study suggests the effect of Observational Conditioning and Stimulus Enhancement to last at least 6 minutes instead of 10 minutes in the earlier paper. The sample size for this study includes adults instead of children, is more numerous (44 versus 30), and is virtual in nature with an artificial agent as a demonstrator.

4.7 Limitations and Future Work

The study is limited in nature for two reasons. First, it is very difficult to find other studies in human-human social transmission that can be replicated for Human Agent Interaction where the agent is artificial. Therefore, the scope is limited to suggesting that such types of social transmission are detectable and quantifiable. Second, while original plans to conduct this study were in person, the complete halt of in-person activities during COVID-19 forced this study to be conducted online.

Therefore, these two limitations become scope for future work, i.e. the goal would be to conduct in-person experiments with real robots for human-robot interaction in order to verify the results gained in this remote study, and also to expand the studies to further incorporate other types of social transmission mechanisms. Certain other limitations in this study include unequal sample sizes regarding gender and gaming experience, which again was due to limitations of using crowd-sourcing methods, however, taken care of by the chosen statistical tests.

While two experimental conditions were studied here, it would also be desirable to compare these two conditions with a neutral condition. While we drew inspiration from [35] who compared only the positive and negative conditions, others [53] utilize the approach with a neutral condition as well. Finally, variations of the virtual environment, tasks, and the nature of the demonstrator (e.g. whether a robotic or human-like or animal-like agent) could be studied further. All of these factors make for exciting possibilities to conduct future work.

4.8 Conclusion

An online study was designed to understand if simpler forms of social learning (i.e. simpler than imitation) can be observed between artificial agents and human participants being present in a virtual, game-like experimental environment. Results from objective and subjective data collected during the online game, which was carefully designed to perform stimulus pairing, point towards successful social transmission between a robotic virtual agent and human participants of information utilizing a mixture of methods, as identified in the literature as Observational Conditioning, Stimulus Enhancement and possibly Response Facilitation. The study closely emulates work done previously in the form of mother-child interaction, [35], and to some extent, human-animal interaction [53]. Neither gender nor previous gaming experience seem to play any significant role in the efficacy of social transmission of information in our study.

The study, to the best of our knowledge, is a novel approach in the field of Human Interaction with Artificial Agents, inspired by, and trying to replicate, as much as possible, experiments in behavioural sciences. Further, studies with in-person participation and real robots would be beneficial, once such research is possible, to verify and extend the results.

Chapter 5

Conclusion

L'homme n'est rien, l'oeuvre – tout, translated to 'The man is nothing, the work — all'.

Gustave Flaubert's letter to George Sand, Dec. 1875

For the concluding phase of this thesis, it is good to refer back to Figure 1.1 and go over the objectives once again. The intended goals of the thesis were to explore social learning in multi-agent robotics. The first study used robots as the target of social transmission. In the second study, a human participant was the target of the social transmission with the robot reacting in a certain way to a stimulus. For study 1, we utilized Local Enhancement and for the second, a combination of Stimulus Enhancement and Observational Conditioning were, according to the literature, the most likely candidates of social transmission occurring.

5.1 Summary of Findings

Here, we have a second look at the research questions and hypotheses of the two studies with brief answers to how those were addressed in this thesis.

5.1.1 Study 1

The research questions (RQs) were defined as follows:

RQ1 : How does the performance of social exploration in robots compare to asocial exploration under time constraints?
with the Hypothesis
H1 : Based on social learning theory, a simple social learning mechanism, when infused in a multi-agent robot society, should perform better than simple asocial explorers.

While we did find significant differences between the reward cache locations discovered for asocial exploration and social learning, the answer is not simply that social learning performs much better. Because we determined that social learning, in our case LE, is a feature that can be switched on or off, there are various degrees of social learning and for what was defined as critical and conditional learning, the number of reward cache locations foraged was higher on average for most cases, but not significantly so.

It is also worth questioning how much better social learning does when compared to asocial exploration. Tables 3.2 and 3.6 show the means and significant differences between social and asocial exploration. Figure 3.6 then shows the rate of return of robotic agents to one of the reward cache locations in order to get more food / fuel and continue their search for more reward locations. First, at lower reward densities i.e. 5 and 10, social learning produces 2.067 and 2.667 rewards on average respectively, while spiral asocial exploration produces 1 and 2.2 reward locations respectively according to 3.2. The percentage of reward caches located on average by an agent is therefore 41% and 53% respectively for social learning, and 20% and 44% respectively for asocial exploration.
Specific to social learning then, 3.6 shows that for those low reward density situations, robots are then able to reach a reward cache location within 400 iterations 36.67% and 43.3% of times. While LE tests for asocial exploration were not performed because LE does not make sense in an asocial context, we can see that social exploration does make a large difference for 'harsher' environments where cache locations are not as widely available.

We are therefore able to confirm Hypothesis 1.

RQ2 : How does the performance of group exploration in social and asocial exploration scenarios compare under conditions of varying amount of rewards?

Group exploration under varying reward densities was of great interest to us since at lower reward densities, groups of primates behave very differently from when they realize that the environment is rich with rewards / food. For the groups of robots investigated here, social learners performed much better, sometimes significantly better, at higher reward densities (20, 30, 40 and 50). For lower reward densities, social learners performed better than asocial explorers, and while the differences were not significant, they are enough to let social explorers survive for longer. Table 3.2 gives the mean number of locations that were foraged / discovered for these experiments.

One point of interest for RQ's 1 and 2 is the consistent observation that critical and conditional learning do not discover locations at a significantly higher rate than asocial explorers. This is further confirmed for RQ3 as well.

RQ3 : What are the effects of learning socially first (critical learning) as compared to learning socially later (conditional learning)?

Hypothesis 2 (H2) : Critical social learning should perform better than conditional social learning in a stable environment (where reward locations do not change often) when primary reward density is low. This follows from the discussion in Chapter 2.3.2 and from earlier studies conducted by [60] and [57].

Comparing the mean number of rewards foraged, neither type of social learning seems to have outperformed the other, and the mean values for both are the same. Here, H2 is also rejected, i.e. critical learning does not perform any better than conditional learning. We believe this is due to the reason that the environment is stable and for reward caches, neither the locations nor the values of the cache change. While [60] is able to provide information on what happens when agents under social learning explore unstable environments, they do not explicitly talk of either critical or conditional learning in unstable environments. This is therefore an important avenue for future investigation.

5.1.2 Study 2

The research questions asked for study 2 were:

RQ1: Does the perceived reaction (positive, negative) of the demonstrator (robot) affect the observer's (human participant's) response? If so, to what extent?

H1 for RQ1: Distance of the participant to the animal in the positive condition should be lower than in the negative condition. We expect participants in the positive condition to have a more positive attitude towards the alien animal. For example, in [35], children observe and infer the nature of the relationship between their carer and the novel creature by recognizing disgust or fear, and replicating the same avoidance behavior towards these creatures.

Distance measures were taken at instances of time, i.e. at 15s, 30s, 45s, 60s and 100s. Two calculations were made with these distances. *Absolute* distance measures the exact euclidean distance of the participant to the animal at that point in time. *Mean* distance is the calculation of mean distances of the participant *up to* that point in time. Both measures show that there are significant differences between participants and the animal.

For the mean distance measure, which is considered more reliable due to the fact that it averages the distances up until a point in time, distances at 30s, 45s and 60s as shown in Figure 4.5a are significantly different. We believe the other distances are not significant because at 15s, participants are still deciding how to react, or have begun reacting in a specific manner but the differences are not significant, and at 100s, the participants either become curious towards the animal, or simply go back to searching for the resources, which is their primary objective. It is to be noted that the distance measure in the Figure at 100s is lower for the blue bar (positive condition) and higher for the orange bar (negative condition), which means the effect can still be seen, but it is just not significant.

The absolute distance measure is considered less reliable because regardless of the participants' reactions, they might at that instant be either too close or further away from the animal. Regardless, it is still an objective measure, and shows significant differences at 15s and 30s. While absolute distance at both 45s and 60s show major differences, and the blue bar is still lower than the orange bar, the differences are not significant.

We can therefore say that the objective data points towards social learning taking place in-game.

H2 for RQ1: Participants in the positive condition should have a more positive and less violent perception of the creature as compared to the negative condition.

Chapter 4.5.3 provides more details about the significance tests run on the perceptions of participants towards the animal in the two conditions using two questions (listed in Table 4.3). In short, no significant differences were found between the negative and positive condition, which means Hypothesis 2 was not confirmed.

RQ2: Does the participant's gaming experience or gender affect their responses towards the creature? If so, to what extent?

No supporting evidence was found for any differences between the two genders (as self-identified by participants) in the positive or negative conditions, nor was there any proof that their previous gaming experiences were a factor in their decision making. While the differences might be there, we believe the number of participants required to observe this effect is too low in the current study.

5.2 Contributions

5.2.1 Study 1

The goals and contributions of this study were to establish social exploration is more optimal manner of exploration in a foraging scenario where a group of robots might depend on each other to identify locations of reward caches that help them survive.

One of the other goals of the study was to try to accomplish autonomy for robotic agents using Local Enhancement. The autonomy would be in the context of agents finding reward caches at a location (either asocially or socially) that hold multiple rewards (hence the word cache) which they would consume one at a time as and how they were required to keep themselves powered up.

Statistical analysis of robot exploration of the two types shows significant differences at larger reward density scenarios. While differences between the two types are not significant for lower reward densities, these differences are enough to make a large contribution and keep the group of robots working for longer and to discover more reward locations.

We envisage the applications of the concepts and results of this study to help search over large areas that are unmapped and unknown by multi-agent robotic systems. The multi-agent system may or may not act socially and would require a degree of autonomy to search for either constrained or low occurring resources, or specific targets in large areas.

5.2.2 Study 2

In general, with the conclusions made from Study 2, objective seems to support the hypothesis that social learning does occur when participants learn to associate stimuli from an experienced demonstrator. There is insufficient evidence to say the same with subjective data, but we believe the effect is present, just not visible with the current number of participants.

However, since objective data is considered stronger due to it's nature, we do strongly believe that this study is the first to show social transmission of information between a virtual robotic agent and human participants with observable effects. Further, this study establishes certain objective measures through which these effects can be calculated. The discussion in Chapter 4.6.4 also lends credence to this, and we utilize formal definitions from the literature, specifically from [43] and [41] to establish indications of social learning happening in the conditions that were described.

5.2.3 Validity of online studies

Since the platform used for conducting Study 2 is entirely virtual (a game environment), it is worth asking whether this study can be emulated in real world conditions, and how close the study gets to evoking actual emotional responses from participants. The pertinent practical question being, can we compare emotional arousal in games to experiencing emotions in the real world?

While there are different opinions about the extent of the human reaction to events in digital media (including games and movies), [63] pointed out that the responses to media events are remarkably similar to real world responses. While there is some agreement that the immediate emotional response is very similar, the longer term effects might be more blunted for virtual events. [15] asserts that previous experiences do play a strong role in evoking emotional responses, especially from the real world, however in the virtual world these are less intense than the same happening in the real world. [15] and [50] argue that previous experiences may allow participants to cope with fear based scenarios more effectively and allow a more balanced reaction towards these events.

Given the current trend of artificial agents playing increasingly bigger roles in human society, understanding how these agents may influence humans is an important question to consider.

5.3 Limitations and Future Work

Both of these studies are exploratory in nature, and therefore some important work remains to be completed. Both studies were meant to explore concepts further using physical robots, but had to be continued in an online mode due to restrictions from COVID-19.

Study 1

Firstly, Study 1 relies extensively on concepts of embodiment, i.e. the capacity of a robot to individually sense reward cache locations and, if in the social learning mode, to transmit or receive this information. How this can be done is an important detail that needs to be resolved. While the study was conducted in an online manner, a real time implementation would allow us to engineer details that allow local enhancement to actually take place in real robots.

Second, the study was limited to 3 robots to demonstrate certain concepts. An important future contribution is using more than 3 robots to establish rules for more realistic number in a group of robots.

Third, it is important to investigate the advantages and disadvantages of different types of social learning, specifically critical and conditional learning in an unstable environment. Unstable environments mean a changing density of reward locations, which is more realistic. It also means reward caches can be exploited only to a certain extent and get exhausted later. Under these conditions, the hypothesis would be that critical learning turns out to be better.

Fourth, very simple rules that constitute other types of social learning, such as 'Copy when uncertain' or 'Copy when a conspecific has higher success rate' can also be implemented and tested on the above conditions.

Lastly, to tie all these ideas up, because the study aspires to look at foraging optimally *as a society*, and not individually, game theory is a very promising avenue of investigation. Currently, an individual agent shares locations only when it is programmed to do so. However, ideally we would like for the robotic agent to have autonomy in decision making on when to share reward cache locations in a cooperative manner while conserving its own battery life. Since there are several ways of achieving a Nash Equilibrium, some approaches tend to settle on a strategy that leads to an Evolutionarily Stable Strategy for maximal reward as a group [25] [73] , whereas others take it as cooperative [73] or non-cooperative game.

Study 2

An important future work is to extend the Positive and Negative conditions to include a Neutral condition. This is because while Gerull and Rapee in [35] use a positive and negative condition, some original work on Observational Conditioning by Mineka et al. in [53] use the neutral condition as well.

Further, a comparison between a human demonstrator and the robotic demonstrator already present will be able to tell us if there are any significant differences due to the embodiment of the demonstrator.

Lastly, an additional condition that explores the experience of the demonstrator and whether this factor has a difference on the effect of observational conditioning is also planned.

All this extension needs to be conducted in a real-world environment, and conversion of the scenario as described for Study 2 into something more realistic will have to be done to create a stimulus. This would be one of the bigger challenges. We believe due to the evidence presented from earlier studies in Chapter 5.2.3 that the metrics for evaluation of social learning should match the ones we have found in these studies as well.

References

[1] A. Acerbi and S. Nolfi. Social learning and cultural evolution in embodied and situated agents. In *2007 IEEE Symposium on Artificial Life*, pages 333–340, 2007.

[2] Ran Aisner and Joseph Terkel. Ontogeny of pine cone opening behaviour in the black rat, rattus rattus. *Animal Behaviour*, 44:327–336, 1992.

[3] Michael A Arbib. Schema theory. *The Encyclopedia of Artificial Intelligence*, 2:1427–1443, 1992.

[4] Michael A Arbib, Péter Érdi, János Szentágothai, Alice Szentagothai, et al. *Neural organization: Structure, function, and dynamics*. MIT press, 1998.

[5] Willem A Arrindell, Paul MG Emmelkamp, et al. Phobic dimensions: Reliability and generalizability across samples, gender and nations: The fear survey schedule (fss-iii) and the fear questionnaire (fq). *Advances in Behaviour Research and Therapy*, 6(4):207–253, 1984.

[6] Ida Bailey, Julia P Myatt, and Alan M Wilson. Group hunting within the carnivora: physiological, cognitive and environmental influences on strategy and cooperation. *Behavioral Ecology and Sociobiology*, 67(1):1–17, 2013.

[7] G Bennett and JR Galef. Social transmission of acquired behavior: a discussion of tradition and social learning in vertebrates. *Advances in the Study of Behavior*, page 77, 1976.

[8] Aude Billard, Sylvain Calinon, Ruediger Dillmann, and Stefan Schaal. Robot programming by demonstration. *Springer handbook of robotics*, pages 1371–1394, 2008.

[9] Richard W. Byrne. The evolution of intelligence. In *Behavior and Evolution*, pages 223–265. Cambridge University Press, New York, NY, US, 1994.

[10] Richard W Byrne. Imitation without intentionality. using string parsing to copy the organization of behaviour. *Animal Cognition*, 2(2):63–72, 1999.

[11] Richard W Byrne. Imitation of novel complex actions: What does the evidence from animals mean? *Advances in the Study of Behavior*, 31:77–105, 2002.

[12] Richard W Byrne and Joanne E Tanner. Gestural imitation by a gorilla: evidence and nature of the capacity. *International Journal of Psychology and Psychological Therapy*, 6(2):215–231, 2006.

[13] Richard W Byrne and Michael Tomasello. Do rats ape? *Animal Behaviour*, 50(5):1417–1420, 1995.

[14] Maya Cakmak, Nick DePalma, Rosa I Arriaga, and Andrea L Thomaz. Exploiting social partners in robot learning. *Autonomous Robots*, 29(3-4):309–329, 2010.

[15] Joanne Cantor. Fright reactions to mass media. In *Media effects*, pages 303–319. Routledge, 2009.

[16] TM Caro. Effects of the mother, object play, and adult experience on predation in cats. *Behavioral and Neural Biology*, 29(1):29–51, 1980.

[17] Luigi Cattaneo and Giacomo Rizzolatti. The mirror neuron system. *Archives of neurology*, 66(5):557–560, 2009.

[18] William G Cochran. Some methods for strengthening the common χ 2 tests. *Biometrics*, 10(4):417–451, 1954.

[19] Michael Cook, Susan Mineka, Bonnie Wolkenstein, and Karen Laitsch. Observational conditioning of snake fear in unrelated rhesus monkeys. *Journal of abnormal psychology*, 94(4):591, 1985.

[20] Colin Das, Aaron Becker, and Timothy Bretl. Probably approximately correct coverage for robots with uncertainty. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1160–1166. IEEE, 2011.

[21] Kerstin Dautenhahn and Chrystopher L Nehaniv. *Imitation in Animals and Artifacts*. Complex Adaptive Systems. MIT Press, Cambridge, Massachusetts, 2002.

[22] Wim De Mulder, Steven Bethard, and Marie-Francine Moens. A survey on the application of recurrent neural networks to statistical language modeling. *Computer Speech & Language*, 30(1):61–98, 2015.

[23] Yan Duan, Marcin Andrychowicz, Bradly Stadie, OpenAI Jonathan Ho, Jonas Schneider, Ilya Sutskever, Pieter Abbeel, and Wojciech Zaremba. One-shot imitation learning. In *Advances in neural information processing systems*, pages 1087–1098, 2017.

[24] Magnus Enquist, Kimmo Eriksson, and Stefano Ghirlanda. Critical social learning: a solution to rogers's paradox of nonadaptive culture. *American Anthropologist*, 109(4):727–734, 2007.

[25] Ahmad Esmaeili, Zahra Ghorrati, and Eric Matson. Multi-agent cooperation using snow-drift evolutionary game model: Case study in foraging task. In *2018 Second IEEE International Conference on Robotic Computing (IRC)*, pages 308–312. IEEE, 2018.

[26] Dorothy M Fragaszy and Susan Perry. *The biology of traditions: models and evidence.* Cambridge University Press, 2008.

[27] G Matthew Fricke, Joshua P Hecker, Antonio D Griego, Linh T Tran, and Melanie E Moses. A distributed deterministic spiral search algorithm for swarms. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4430–4436. IEEE, 2016.

[28] Johannes Fritz, Agnes Bisenberger, and Kurt Kotrschal. Stimulus enhancement in greylag geese: socially mediated learning of an operant task. *Animal Behaviour*, 59(6):1119–1125, 2000.

[29] Ryusuke Fujisawa, Shigeto Dobata, Ken Sugawara, and Fumitoshi Matsuno. Designing pheromone communication in swarm robotics: Group foraging behavior mediated by chemical substance. *Swarm Intelligence*, 8(3):227–246, 2014.

[30] Hamdy Mohamed Abou Gabal and Mohamed Abd Allah El-Hadidy. Optimal searching for a randomly located target in a bounded known region. *International Journal of Computing Science and Mathematics*, 6(4):392–403, 2015.

[31] Bennett G Galef. The question of animal culture. *Human nature*, 3(2):157–178, 1992.

[32] Bennett G Galef Jr. Imitation in animals: History, definition, and interpretation of data from the psychological laboratory. In *Social learning*, pages 15–40. Psychology Press, 2013.

[33] Bennett G Galef Jr and David J White. Mate-choice copying in japanese quail, coturnix coturnix japonica. *Animal Behaviour*, 55(3):545–552, 1998.

[34] Simon Garnier, Faben Tache, Maud Combe, Anne Grimal, and Guy Theraulaz. Alice in pheromone land: An experimental setup for the study of ant-like robots. In *2007 IEEE Swarm Intelligence Symposium*, pages 37–44. IEEE, 2007.

[35] Friederike C Gerull and Ronald M Rapee. Mother knows best: effects of maternal modelling on the acquisition of fear and avoidance behaviour in toddlers. *Behaviour research and therapy*, 40(3):279–287, 2002.

[36] L Lefebvreand L Giraldeau and L Lefebvre. Is social learning an adaptive specialization. *Social learning in animals: the roots of culture (C. M Heyesand B. G Galef Jr eds.). Academic Press, San Diego, California*, pages 107–128, 1996.

[37] Luc-Alain Giraldeau, Thomas J Valone, and Jennifer J Templeton. Potential disadvantages of using socially acquired information. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1427):1559–1566, 2002.

[38] Jane Goodall. The chimpanzees of gombe: Patterns of behavior. *Cambridge Mass*, 1986.

[39] Samuel D Gosling, Peter J Rentfrow, and William B Swann Jr. A very brief measure of the big-five personality domains. *Journal of Research in personality*, 37(6):504–528, 2003.

[40] Cecilia Heyes. Evolution, development and intentional control of imitation. *Philosophical transactions. Biological sciences*, 364(1528):2293–2298, 2009.

[41] Cecilia M Heyes. Social learning in animals: categories and mechanisms. *Biological Reviews*, 69(2):207–231, 1994.

[42] Nicholas Hoff, Robert Wood, and Radhika Nagpal. Distributed colony-level algorithm switching for robot swarm foraging. In *Distributed Autonomous Robotic Systems*, pages 417–430. Springer, 2013.

[43] William Hoppitt and Kevin N Laland. *Social learning: an introduction to mechanisms, methods, and models*. Princeton University Press, 2013.

[44] Ahmed Hussein, Mohamed Medhat Gaber, Eyad Elyan, and Chrisina Jayne. Imitation learning: A survey of learning methods. *ACM Computing Surveys (CSUR)*, 50(2):21, 2017.

[45] Jungsik Hwang and Jun Tani. Seamless integration and coordination of cognitive skills in humanoid robots: A deep learning approach. *IEEE Transactions on Cognitive and Developmental Systems*, 10(2):345–358, 2018.

[46] Michael Krützen, Janet Mann, Michael R Heithaus, Richard C Connor, Lars Bejder, and William B Sherwin. Cultural transmission of tool use in bottlenose dolphins. *Proceedings of the National Academy of Sciences*, 102(25):8939–8943, 2005.

[47] Kevin N Laland. *Darwin's unfinished symphony: How culture made the human mind*. Princeton University Press, 2018.

[48] Kevin N Laland and Bennett G Galef. *The question of animal culture*. Harvard University Press, 2009.

[49] Ellouise Leadbeater and Lars Chittka. The dynamics of social learning in an insect model, the bumblebee (*Bombus terrestris*). *Behavioral Ecology and Sociobiology*, 61(11):1789–1796, 2007.

[50] Teresa Lynch and Nicole Martins. Nothing to fear? an analysis of college students' fear experiences with video games. *Journal of Broadcasting & Electronic Media*, 59(2):298–317, 2015.

[51] Cara C MacInnis, Harrison CD Boss, and Joshua S Bourdage. More evidence of participant misrepresentation on mturk and investigating who misrepresents. *Personality and Individual Differences*, 152:109603, 2020.

[52] O. Michel. Webots: Professional mobile robot simulation. *Journal of Advanced Robotics Systems*, 1(1):39–42, 2004.

[53] Susan Mineka and Michael Cook. Social learning and the acquisition of snake fear in monkeys. In *Social learning*, pages 63–86. Psychology Press, 2013.

[54] Chrystopher L Nehaniv and Kerstin Dautenhahn. Like me?-measures of correspondence and imitation. *Cybernetics & Systems*, 32(1-2):11–51, 2001.

[55] Chrystopher L Nehaniv and Kerstin Dautenhahn, editors. *Imitation in animals and artifacts*. MIT press, 2002.

[56] Chrystopher L Nehaniv and Kerstin Dautenhahn, editors. *Imitation and social learning in robots, humans and animals: behavioural, social and communicative dimensions*. Cambridge University Press, 2007.

[57] Jason Noble and Daniel W Franks. Social learning in a multi-agent system. *Computing and Informatics*, 22(6):561–574, 2012.

[58] Boris Palameta and Louis Lefebvre. The social transmission of a food-finding technique in pigeons: what is learned? *Animal behaviour*, 33(3):892–896, 1985.

[59] Daniel Pickem, Paul Glotfelter, Li Wang, Mark Mote, Aaron Ames, Eric Feron, and Magnus Egerstedt. The robotarium: A remotely accessible swarm robotics research testbed. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1699–1706. IEEE, 2017.

[60] Lenka Pitonakova, Richard Crowder, and Seth Bullock. Information flow principles for plasticity in foraging robot swarms. *Swarm Intelligence*, 10(1):33–63, 2016.

[61] Harish Ravichandar, Athanasios S Polydoros, Sonia Chernova, and Aude Billard. Recent advances in robot learning from demonstration. *Annual Review of Control, Robotics, and Autonomous Systems*, 3:297–330, 2020.

[62] Simon M Reader, Jeremy R Kendal, and Kevin N Laland. Social learning of foraging sites and escape routes in wild trinidadian guppies. *Animal Behaviour*, 66(4):729–739, 2003.

[63] Byron Reeves and Clifford Nass. *The media equation: How people treat computers, television, and new media like real people*. Cambridge university press Cambridge, UK, 1996.

[64] Alan R Rogers. Does biology constrain culture? *American Anthropologist*, 90(4):819–831, 1988.

[65] Joe Saunders, Chrystopher L Nehaniv, and Kerstin Dautenhahn. Teaching robots by moulding behavior and scaffolding the environment. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 118–125. ACM, 2006.

[66] Thomas Schmickl and Heiko Hamann. Beeclust: A swarm algorithm derived from honeybees. *Bio-inspired computing and communication networks*, pages 95–137, 2011.

[67] Madeleine Y Stepper, Cathleen M Moore, Bettina Rolke, and Elisabeth Hein. Illusory size determines the perception of ambiguous apparent motion. *Psychonomic Bulletin & Review*, 27(6):1230–1238, 2020.

[68] Andrea L Thomaz and Maya Cakmak. Social learning mechanisms for robots. Georgia Institute of Technology, 2009.

[69] Edward Thorndike. *Animal intelligence: Experimental studies.* New York, The Macmillan company, 1911.

[70] William Homan Thorpe. Learning and instinct in animals. 1956.

[71] Michael Tomasello. *Social learning in animals : the roots of culture.* Academic Press, San Diego, 1996.

[72] Michael Tomasello. *The cultural origins of human cognition.* Harvard university press, 2009.

[73] Yuehai Wang, Jie Liu, and Wei Meng. Cooperative algorithm for multi-agent foraging task based on modified hawk-dove game. In *2007 International Conference on Computational Intelligence and Security Workshops (CISW 2007)*, pages 179–182, 2007.

[74] Andrew Whiten. The second inheritance system of chimpanzees and humans. *Nature*, 437(7055):52–55, 2005.

[75] Andrew Whiten, Jane Goodall, William C McGrew, Toshisada Nishida, Vernon Reynolds, Yukimaru Sugiyama, Caroline EG Tutin, Richard W Wrangham, and Christophe Boesch. Cultures in chimpanzees. *Nature*, 399(6737):682–685, 1999.

[76] Yuichi Yamashita and Jun Tani. Emergence of functional hierarchy in a multiple timescale neural network model: a humanoid robot experiment. *PLoS computational biology*, 4(11):e1000220, 2008.

[77] Tianhe Yu, Chelsea Finn, Annie Xie, Sudeep Dasari, Tianhao Zhang, Pieter Abbeel, and Sergey Levine. One-shot imitation from observing humans via domain-adaptive meta-learning. *arXiv preprint arXiv:1802.01557*, 2018.

[78] Thomas R Zentall. An analysis of imitative learning in animals. In Bennett G. Galef Jr. Cecilia M. Heyes, editor, *Social Learning In Animals: The Roots of Culture*, pages 221–243. ACADEMIC PRESS, INC., 24-28 Oval Road, London, NW1 7DX, United Kingdom, 1996.

[79] Thomas R Zentall. Imitation in animals: evidence, function, and mechanisms. *Cybernetics & Systems*, 32(1-2):53–96, 2001.

[80] Fuzhen Zhuang, Zhiyuan Qi, Keyu Duan, Dongbo Xi, Yongchun Zhu, Hengshu Zhu, Hui Xiong, and Qing He. A comprehensive survey on transfer learning. *arXiv preprint arXiv:1911.02685*, 2019.

APPENDICES

Appendix A

Ethics Clearance

UNIVERSITY OF WATERLOO

Notification of Ethics Clearance to Conduct Research with Human Participants

---

Principal Investigator: Kerstin Dautenhahn (Electrical and Computer Engineering)

Collaborator: Shruti Chandra (Electrical and Computer Engineering)

Student investigator: Owais Hamid (Systems Design Engineering)

Co-Investigator: Chrystopher Nehaniv (Systems Design Engineering)

File #: 42864

Title: Forest Foraging - Experiences in a Virtual World

The Human Research Ethics Committee is pleased to inform you this study has been reviewed and given ethics clearance.

Initial Approval Date: 03/09/21 (m/d/y)

University of Waterloo Research Ethics Committees are composed in accordance with, and carry out their functions and operate in a manner consistent with, the institution's guidelines for research with human participants, the Tri-Council Policy Statement for the Ethical Conduct for Research Involving Humans (TCPS, 2nd edition), International Conference on Harmonization: Good Clinical Practice (ICH-GCP), the Ontario Personal Health Information Protection Act (PHIPA), the applicable laws and regulations of the province of Ontario. Both Committees are registered with the U.S. Department of Health and Human Services under the Federal Wide Assurance, FWA00021410, and IRB registration number IRB00002419 (HREC) and IRB00007409 (CREC).

This study is to be conducted in accordance with the submitted application and the most recently approved versions of all supporting materials.

Expiry Date: 03/10/22 (m/d/y)

Multi-year research must be renewed at least once every 12 months unless a more frequent review has otherwise been specified. Studies will only be renewed if the renewal report is received and approved before the expiry date. Failure to submit renewal reports will result in the investigators being notified ethics clearance has been suspended and Research Finance being notified the ethics clearance is no longer valid.

Level of review: Delegated Review

Signed on behalf of the Human Research Ethics Committee

Heather Root, Senior Manager, Ethics, heather.root@uwaterloo.ca, 519-888-4567, ext. 30469

This above named study is to be conducted in accordance with the submitted application and the most recently approved versions of all supporting materials.

Documents reviewed and received ethics clearance for use in the study and/or received for information:

file: HITInformation_v1_42864.pdf

file: StudyProcedures_v2_42864.docx.pdf

file: InformationandConsentForm_v2_42864.docx.pdf

file: EndMessage_v1_42864.pdf

file: Debriefing_v2_42864.docx.pdf

Approved Protocol Version 2 in Research Ethics System

This is an official document. Retain for your files.

You are responsible for obtaining any additional institutional approvals that might be required to complete this study.

Appendix B

Game Experience

As described earlier, the game experience for all participants was divided into three parts, the pre-game questionnaire, the game itself served through WebGL, and the post-game questionnaire.

B.1 Pre-Game Procedure

Your responsibilities as a participant

Participation in this study will approximately take 45-60 minutes. Your answer to the Questionnaire will be logged. This does not require you to download anything onto your computer. The game can be paused and this could increase the time needed for completing the study. The above estimation is based on your continuous engagement with the study.

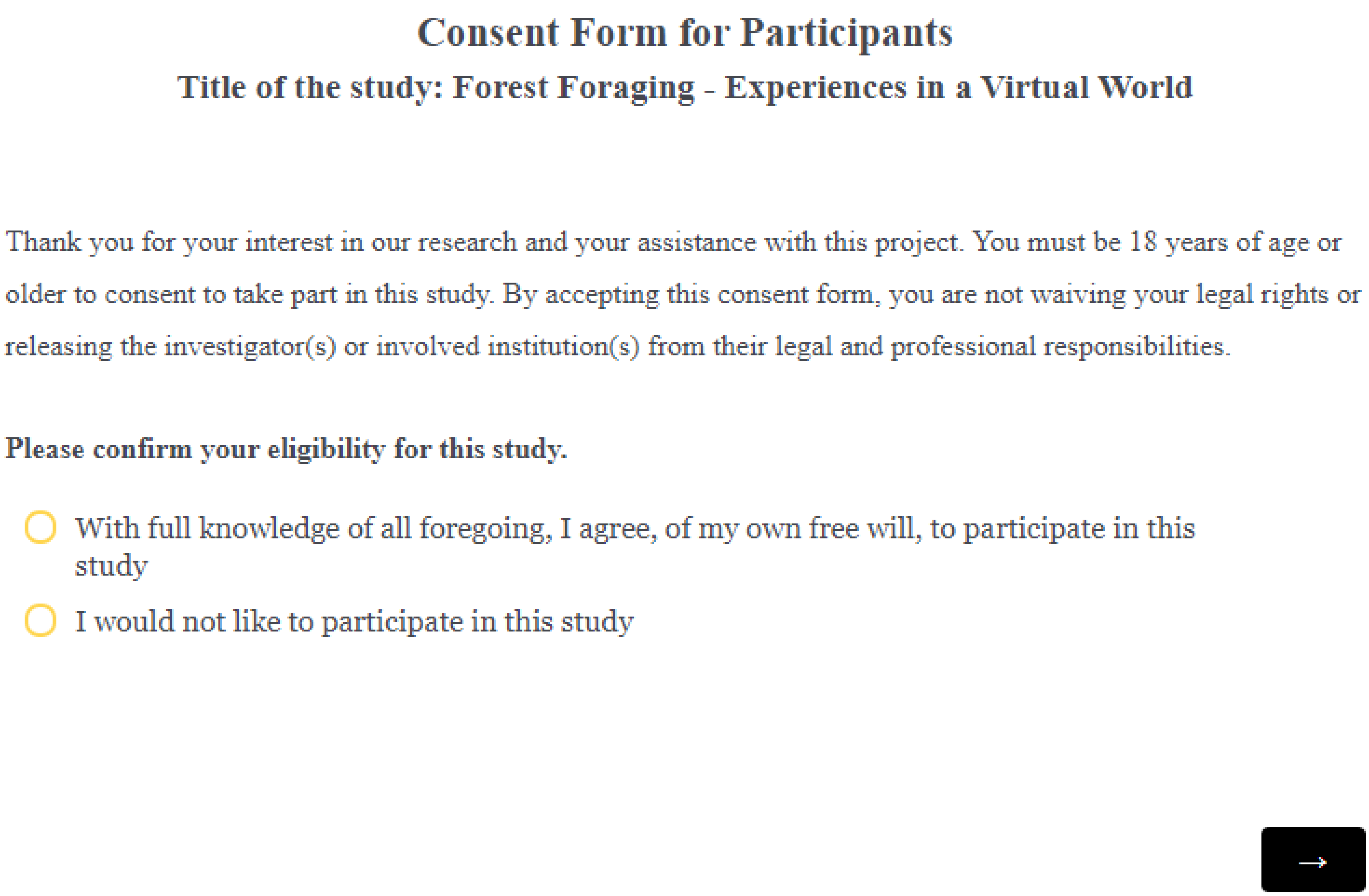
1. The study will start after you give consent by checking the box and pressing the button at the end of this document.
2. First, we will ask your demographic information (only your gender, sex, age, and educational level) and personality. If you do not wish to answer certain questions, you may choose not to answer.
3. In the next step, to make you familiar with the context, you will be able to practice using the interface used for this study prior to the beginning of the experiment.
4. You will then play the game, all on your browser, using a mouse and keyboard. **We highly recommend you use Mozilla Firefox,** not Google Chrome or Microsoft Edge for best performance.
5. You will then finish the post-game questionnaires which include a personality questionnaire and questions on robot personality.
6. At the end, **we will provide you with a code to submit the HIT** and receive your remuneration.

You can complete this study only once. Your responses are valuable information that can help us in our research. Please pay attention and provide the answers as accurately as you can.

Your rights as a participant

Your participation in this study is voluntary. You are free to decline to answer any particular question you do not wish to answer. This does not mean you need to quit the HIT. You may also decide to stop any time. If this is the case, please close the page and leave this HIT. Contact us by email (sirrl.waterloo@gmail.com) with your Mechanical Turk ID in this situation, so we can remunerate you for the proportion of what you completed. In this case your data will be discarded and not be used in the study.

Figure B.1: Rights and Responsibilities of the Participant



Consent Form for Participants

Title of the study: Forest Foraging - Experiences in a Virtual World

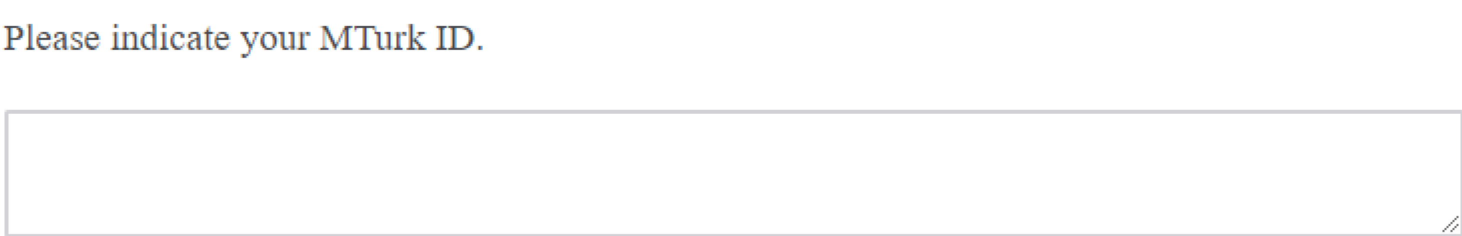
Thank you for your interest in our research and your assistance with this project. You must be 18 years of age or older to consent to take part in this study. By accepting this consent form, you are not waiving your legal rights or releasing the investigator(s) or involved institution(s) from their legal and professional responsibilities.

Please confirm your eligibility for this study.

- With full knowledge of all foregoing, I agree, of my own free will, to participate in this study
- I would not like to participate in this study

→

Figure B.2: Consent Form



Please indicate your MTurk ID.

Figure B.3: Entering the MTurk ID of the participant

HIT Information

Title of the study: Forest Foraging - Experiences in a Virtual World

This is a study that explores how people interact with other agents in a virtual (game) world. You will be playing a game that is set in a forest, which contains some other agents as well.

This game will automatically load on your web browser, and you do not need to download anything. **We recommend you use Mozilla Firefox,** not Google Chrome or Microsoft Edge for best performance. **You simply need a computer/ laptop with a keyboard and mouse**. Your objective is to gather as many energy resources as possible within the given time frame. The game will require about 25 minutes of your time. Please note that the game is time sensitive and requires you to complete certain tasks in time.

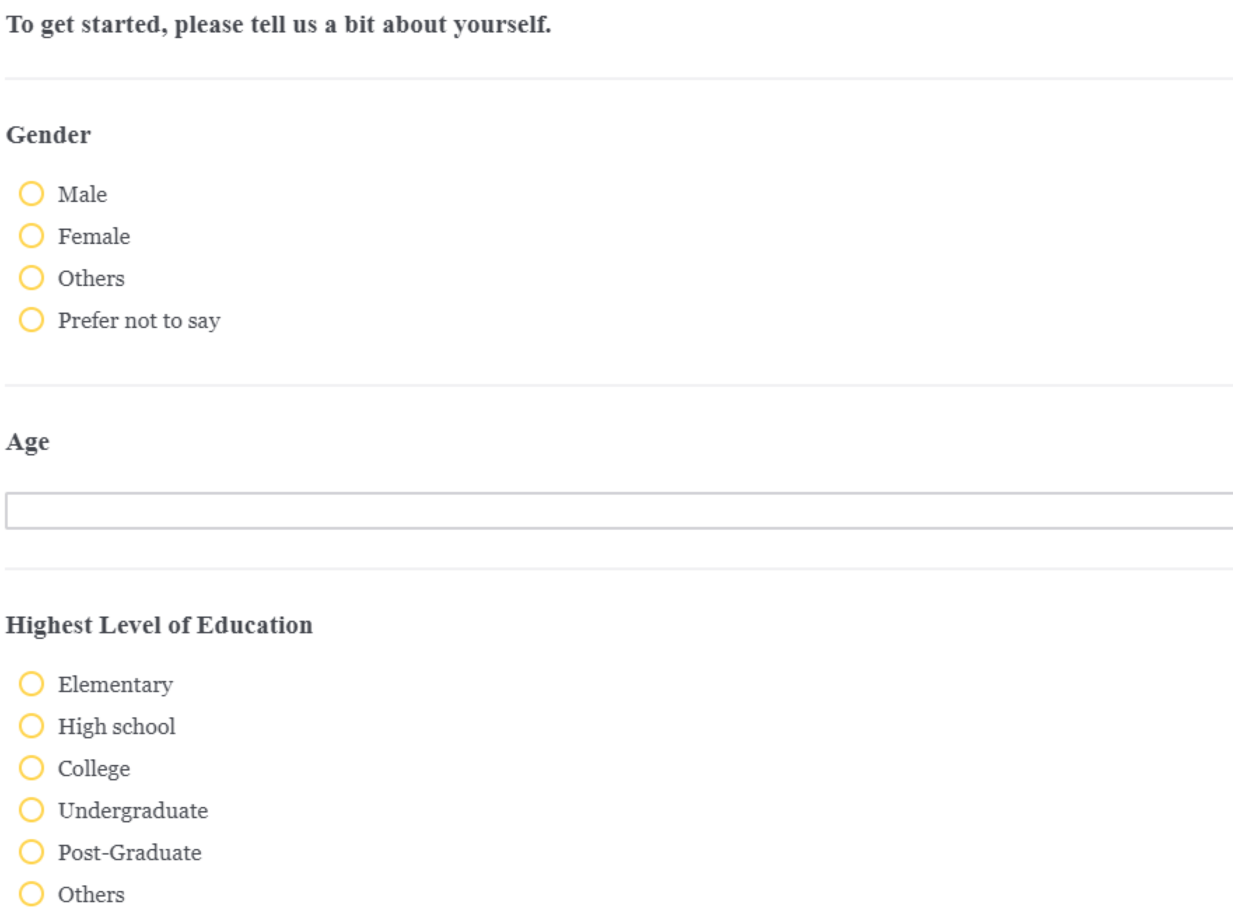
There are also two questionnaires, one before and one after the game. They will require at most 10 minutes.

At the end of the study, you will see a confirmation page with a code. You can close the study and submit this HIT by entering the code on this page. **This study might take a total of 35-45 minutes.** The above estimation is based on your continuous engagement. **You will receive $4 for completing all the steps** (the code will be given to you on the last page of the study), or a pro-rated amount if you wish to stop earlier. In this case, contact us to receive a code, and you will get the $1 base payment, and based on the proportion of the Questions you have answered, up to $4.

Please note the following:

- We will not be able to accept your HIT if you do not provide a valid code.
- You may only join this HIT once.
- It is important to pay attention to the HIT and provide accurate responses.

Figure B.4: Game Information that was given to the participants



To get started, please tell us a bit about yourself.

Gender

- Male
- Female
- Others
- Prefer not to say

Age

Highest Level of Education

- Elementary
- High school
- College
- Undergraduate
- Post-Graduate
- Others

Figure B.5: Pre-game questionnaire

Please choose an option next to each statement to indicate the extent to which you agree or disagree with that statement. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.

	Strongly Disagree	Moderately Disagree	Disagree a Little	Neither Agree nor Disagree	Agree a Little	Agree Moderately	Agree Strongly
I see myself as Extraverted, enthusiastic	○	○	○	○	○	○	○
I see myself as Critical, quarrelsome	○	○	○	○	○	○	○
I see myself as Dependable, self-disciplined	○	○	○	○	○	○	○
I see myself as Anxious, easily upset	○	○	○	○	○	○	○
I see myself as Open to new experiences, complex.	○	○	○	○	○	○	○
I see myself as Reserved, quiet	○	○	○	○	○	○	○
I see myself as Sympathetic, warm.	○	○	○	○	○	○	○
I see myself as Disorganized, careless.	○	○	○	○	○	○	○
I see myself as Calm, emotionally stable.	○	○	○	○	○	○	○
I see myself as Conventional, uncreative.	○	○	○	○	○	○	○

Figure B.6: The Ten Item Personality Index

Game Information

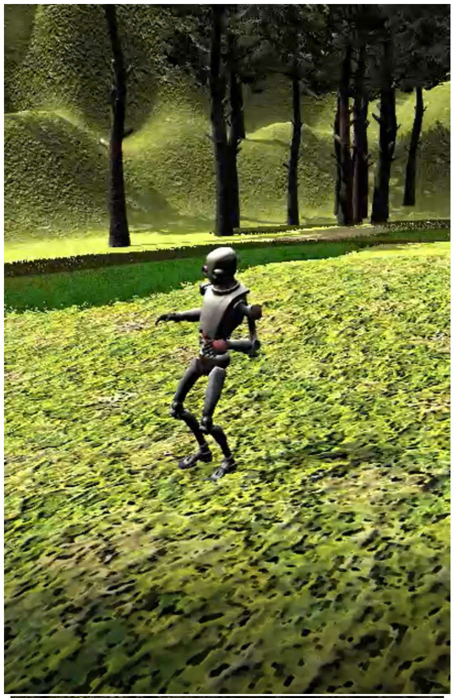
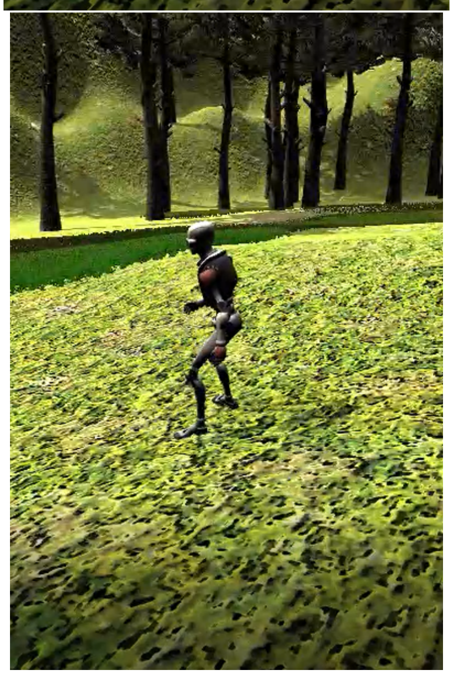
- Note that you would require a computer or laptop, Key and Mouse to play the game.
- The game would require high computational resources and therefore it would ideal **if you do not have any other window or tabs open on your browser**.
- You must **go to the full-screen mode** in order to paly the game properly.
- The game is found to work best with **Mozilla Firefox web browser**. If you find any trouble in playing the game on your browser, please switch to another browser.
- After clicking the game link, you will first enter into the game environment where you would familiarize yourself with the game environment and key controllers.
- Then you will enter into the main game.
- **Instructions on right top corner will always be accessible on the screen**.

Figure B.7: Instructions on how to play the game

B.2 Game-play

The participants were then given the link to the study that took them to the WebGL platform

Figure B.8: The robot can be seen being friendly towards the alien in the positive condition

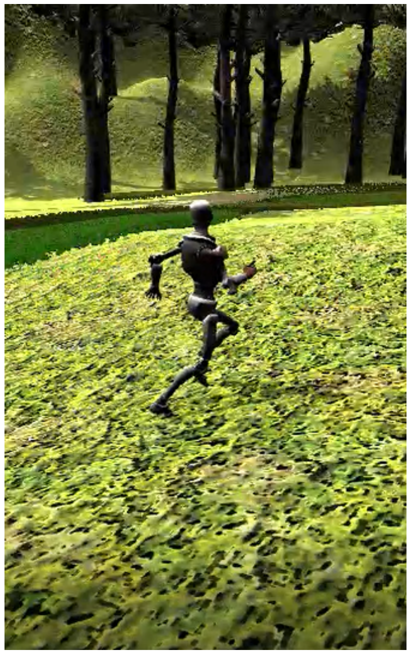


Figure B.10: The reactions of the robot in the negative condition, where the robot turns and runs away in fear

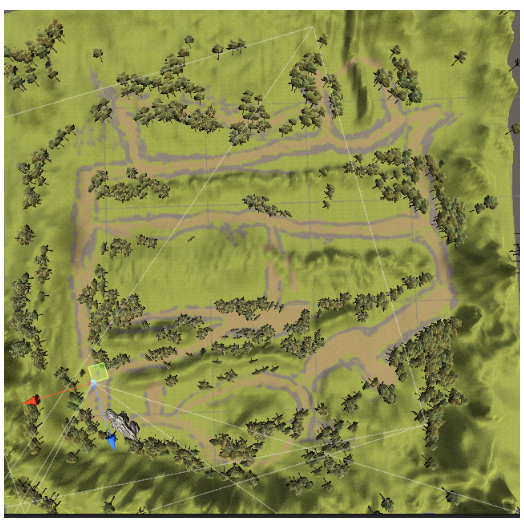


Figure B.11: A bird's eye view of the game world

B.3 Post-game Questionnaire

Both custom-made questions and standard questionnaires were administered during the post-game phase.

Please answer the following questions.

Are you a gamer?

- Yes
- No
- Other

Have you encountered the creature?

- Yes
- No

How does the creature look like?

How did you perceive the behavior of the creature?

When you first saw the creature, what was your reaction?

- I should run away
- I should approach it
- Neither
- Other

Did you encounter the creature a second time?

- ● Yes
- ○ No

What was the behavior of the creature at that time?

Have you approached the creature?

- ○ Yes
- ○ No

How would you behave towards the creature if this were allowed in the game?

What do you think of the creature?

	1	2	3	4	5	
Very frightened	○	○	○	○	○	Not frightened at all

Have you observed the forest inhabitant?

- ○ Yes
- ○ No

What was the reaction of the forest inhabitant towards the creature?

	1	2	3	4	5	
Very frightened	○	○	○	○	○	Not frightened at all
Indifferent	○	○	○	○	○	Not Indifferent at all
Very Curious	○	○	○	○	○	Not curious at all

What was your experience with how the forest inhabitant behaved?

	1	2	3	4	5	
Very Competent	○	○	○	○	○	Not Competent at all
Very Skilled	○	○	○	○	○	Not skilled at all

If you went exploring again, would you trust the forest inhabitant to show you the way to another forest?

	1	2	3	4	5	
Yes, definitely	○	○	○	○	○	Not at all

In your opinion, does the forest inhabitant need more experience in the forest to make better decisions?

	1	2	3	4	5	
Yes, definitely	○	○	○	○	○	Not at all

Which browser did you use to fill the survey and play the game?

- ○ Chrome
- ○ Firefox
- ○ Edge
- ○ Safari
- ○ Other

Did you find any difficulty with the browser to play the game and fill the survey? Please specify.

Please specify any general comments that you would like to share regarding your experience while performing this study.

Figure B.14: Post Game Questionnaire for the Experiments

Please answer the following questions. The questions are presented on a scale of 1 to 5 (1- Not at all, 5- Very much).

	Not at all	A little	A fair amount	Much	Very much
I fear worms	○	○	○	○	○
I fear bats	○	○	○	○	○
I fear flying insects	○	○	○	○	○
I fear crawling insects	○	○	○	○	○
I fear mice	○	○	○	○	○
I fear harmless snakes	○	○	○	○	○
I fear human-like robots	○	○	○	○	○

Figure B.15: The Fear Schedule Survey, specifically the Fear of Harmless Animals Index (ANI)

Warmth	Competence	Discomfort
Organic Social Emotional Compassionate Happy Feeling	Reliable Competent Knowledgeable Interactive Responsive Capable	Awkward Scary Strange Awful Dangerous Aggressive

On a scale from 1 to 7, how well do these words describe the robot you just talked to?

	1 = not at all	2	3	4	5	6	7 = very much so
Dangerous							
Awkward							
Agressive							
Feeling							
Strange							
Knowledgeable							
Reliable							
Happy							
Compassionate							
Awful							
Competent							
Social							
Responsive							
Scary							
Capable							
Emotional							
Interactive							
Organic (Non-mechanical)							

Figure B.16: Robotic Social Attributes Scale

Debriefing Form for Participation in the Research Study: Forest Foraging - Experiences in a Virtual World
University of Waterloo

Thank you for your participation in our study! Your participation is greatly appreciated.

Purpose of the Study

This document is to simply inform you about the purpose of the study, which was not fully explained at the beginning. One of this study's goals is to try to understand your reaction to a new change in the environment and see what effect the reaction of the forest companion has on you. Because we wanted to check your normal reaction, we did not explicitly mention this at the very beginning.

At no point during the game was the creature that came from the Forest a threat to your character.

Confidentiality

You may decide that you do not want your data used in this research. If you would like your data removed from the study and permanently deleted please send an email to sirrl.waterloo@gmail.com with your HIT ID. Whether you agree or do not agree to have your data used for this study, you will still receive compensation for your participation.

Final Report

If you would like to receive a copy of the final report of this study (or a summary of the findings) when it is completed, please feel free to contact us.

Useful Contact Information

If you have any questions or concerns regarding this study, its purpose or procedures, or if you have a research-related problem, please feel free to contact the researcher Owais Hamid at omabdulh@uwaterloo.ca

If you have questions for the Committee contact the Office of Research Ethics, at 1-519-888-4567 ext. 36005 or oreceo@uwaterloo.ca.

Figure B.17: Debriefing the participant regarding the goals / objectives of the game, and thanking them for their participation

Appendix C

The Robotarium

The code described here is a minimal representation of what is required to submit an experiment on the Robotarium, and can also be found on Owais Hamid's Github profile.

```
'''
Owais Hamid, 04/06/20
Robot 1 runs in sawtooth formation
Robot 2 runs in Spiral formation
'''

#Import Robotarium Utilities
import rps.robotarium as robotarium
from rps.utilities.transformations import *
from rps.utilities.graph import *
from rps.utilities.barrier_certificates import *
from rps.utilities.misc import *
from rps.utilities.controllers import *
from math import sin,cos
import numpy as np

# Experiment Constants
iterations = 6000 #Run the simulation/experiment for 5000 steps
↪  (5000*0.033 ~= 2min 45sec)
N=3
```

```
reward1 = 0
reward1_loc = []
reward2 = 0
reward2_loc = []

#Robot 1 waypoint definition. Waypoints define a spiral
x_c = []
y_c = []
x=0
y=0
a=.08
b=.08
angle=0

for i in range(100):
        angle = 0.2*i
        x = (a+b * angle) * cos(angle)                #Needs to be
         ↪  between [-1.5,1.5]
        y = (a+b * angle) * sin(angle)                #Needs to be
         ↪  between [-0.9,0.9]
        x_c = np.append(x_c,x)
        y_c = np.append(y_c,y)
        x_c = np.clip(x_c,-1.5,1.5)
        y_c = np.clip(y_c,-0.9,0.9)

waypoints = np.array([x_c,y_c])                #(2,25)

#Waypoint defining sawtooth formation
percent=30.0
TimePeriod=1.0
Cycles=3
dt=0.12
t=np.arange(-Cycles*TimePeriod/2,Cycles*TimePeriod/2,dt);
pwm= t%TimePeriod<TimePeriod*percent/100
pwm = pwm.astype(float)
pwm = np.where(pwm==1,0.9,-0.9)

waypoints_1 = np.array([t,pwm])                #(2,25)
```

```
close_enough = 0.03; #How close the leader must get to the waypoint to
↪  move to the next one.

#Creating rewards with reproducibility
np.random.seed(0)
reward_x = np.random.uniform(-1.5,1.5,50)
np.random.seed(1)        #Create 50 x_coord for reward
reward_y = np.random.uniform(-0.9,0.9,50)
reward_locs = np.array([reward_x,reward_y])        #Reward Locations

#Initialize states
state = 0
state_1 = 0

#Limit maximum linear speed of any robot
magnitude_limit = 0.15

# For computational/memory reasons, initialize the velocity vector
dxi = np.zeros((2,N))

# Initial Conditions to Avoid Barrier Use in the Beginning.
initial_conditions = np.array([[0,0.5,0.3],[0.5, 0.3, 0.1],[0, 0.2, 0.6]])

# Instantiate the Robotarium object with these parameters
r = robotarium.Robotarium(number_of_robots=N, show_figure=True,
↪  initial_conditions=initial_conditions, sim_in_real_time=True)

# Grab Robotarium tools to do simgle-integrator to unicycle conversions
↪  and collision avoidance
# Single-integrator -> unicycle dynamics mapping
_,uni_to_si_states = create_si_to_uni_mapping()
si_to_uni_dyn = create_si_to_uni_dynamics(angular_velocity_limit=np.pi/2)

# Single-integrator barrier certificates
si_barrier_cert =
↪  create_single_integrator_barrier_certificate_with_boundary()
```

```
# Single-integrator position controller
agent1_controller =
↪  create_si_position_controller(velocity_magnitude_limit=0.15)
agent2_controller =
↪  create_si_position_controller(velocity_magnitude_limit=0.15)

# Plot Graph Connections
x = r.get_poses() # Need robot positions to do this.

r.step()

for t in range(iterations):
        # Get the most recent pose information from the Robotarium. The
         ↪  time delay is ~ 0.033s
        x = r.get_poses()
        xi = uni_to_si_states(x)
        for i in range(1,N):
                # Zero velocities
                dxi[:,[i]]=np.zeros((2,1))

        waypoint = waypoints[:,state].reshape((2,1))
        ws = waypoints_1[:,state_1].reshape((2,1))

        dxi[:,[0]] = agent1_controller(x[:2,[0]], waypoint)
        dxi[:,[1]] = agent2_controller(x[:2,[1]], ws)
        if np.linalg.norm(x[:2,[0]] - waypoint) < close_enough:
                state = (state + 1)%100        #the denominator needs to
                 ↪  be the len(waypoint array)
        if np.linalg.norm(x[:2,[1]] - ws) < close_enough:
                state_1 = (state_1 + 1)%25        #Same with the
                 ↪  denominator here

        #Create reward extension scenario
        for j in range(len(reward_locs[0])):
                if np.linalg.norm(x[:2,[0]] -
                 ↪  reward_locs[:,j].reshape(2,1)) < close_enough:
                        print("Close to",reward_locs[:,j])
                        reward1_loc.append(reward_locs[:,j])
```

```
                #np.delete(reward_locs,j,1)
        if np.linalg.norm(x[:2,[1]] -
        ↪  reward_locs[:,j].reshape(2,1)) < close_enough:
                print("Close to",reward_locs[:,j])
                reward2_loc.append(reward_locs[:,j])
                #np.delete(reward_locs,j,1)

    #Keep single integrator control vectors under specified
    ↪  magnitude
    # Threshold control inputs
    norms = np.linalg.norm(dxi, 2, 0)
    idxs_to_normalize = (norms > magnitude_limit)
    dxi[:, idxs_to_normalize] *=
    ↪  magnitude_limit/norms[idxs_to_normalize]

    #Use barriers and convert single-integrator to unicycle commands
    dxi = si_barrier_cert(dxi, x[:2,:])
    dxu = si_to_uni_dyn(dxi,x)

    # Set the velocities of agents 1,...,N to dxu
    r.set_velocities(np.arange(N), dxu)

    # Iterate the simulation
    r.step()

#Save all reward locations foraged in an npy file. These are retrieved
↪  from the Robotarium after the experiment. Call at end of script to
↪  print debug information and for your script to run on the Robotarium
↪  server properly
reward1_loc = np.unique(reward1_loc,axis=0)
reward2_loc = np.unique(reward2_loc,axis=0)
np.save('reward1.npy',len(reward1_loc))
np.save('reward2.npy',len(reward2_loc))
r.call_at_scripts_end()
```