The Importance of Replication in Uncertain Epistemic Landscapes

Lamar Kiel (6024343)

Supervisor: Dr. Dominik Klein
2nd Examiner: Dr. Benjamin Rin

March 15, 2023

Utrecht University

Acknowledgement

I would like to take this opportunity to thank Dr. Dominik Klein, my supervisor. He has been an amazing help writing this thesis. The discussions we had have been engaging and his insights were incredibly useful for the process of this project. I am sure that the skills obtained and lessons learned will prove to be useful in my future endeavours.

I also would like to thank Michael Weisburg from the University of Pennsylvania and Ryan Muldoon from the University of Buffalo for sharing their model's code on the division of cognitive labour. The model has been a huge inspiration and help for my work.

Abstract

Scientific research is mainly done in groups of scientists working in different parts of a specific research domain. This division of work is called the division of cognitive labour. The division of labour is originally interested in scientists that do different tasks in science. This thesis focuses on different attitudes of scientists and how they cooperate. We present an agent-based model of scientific research in which scientists are divided to explore and exploit unknown areas on an uncertain scientific landscape. The model is a complexification of the models from Muldoon & Weisberg, and Thoma. Scientists aim to find and correctly identify the best approaches for a research problem. However, outcomes of scientific research can be erroneous. Therefore, replication is needed to verify what approaches are best. Failing to replicate is not a direct indicator of false results, but one should pay caution when one cannot reproduce outcomes of scientific research. Three distinct agent types for searching approaches are considered. These types have their unique search rules. The three agents are called: followers, mavericks and replicators. Followers look out for successfully done research and incrementally expand on this knowledge. Mavericks find areas that are not explored already. Lastly, replicators help to correctly identify the best approaches by replicating experiments. As a result, they take away uncertainty. A balance between exploring the landscape and replicating approaches is needed. The model shows that mainly having followers with a low proportion of replicators working on a scientific problem is beneficial. Mavericks are only helpful when scientists are inflexible.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

Philosophy of science is classically interested in characteristics of scientific progress and how to make scientific endeavour successful. This includes questions such as reliability and justified belief. The interest lies in the criteria for evaluating scientific theories. These criteria include predictive success or the potential to solve relevant problems. Though, lately, there also has been a focus on the social structures within science [2, 1, 3]. Examining the scientific structures can be about finding out how to work best collectively. To exploit differences in character and talent and how to use these varieties of character and talent to maximise scientific progress. This topic includes questions regarding how different motivations and personalities can be used together to proceed efficiently in conducting science.

As one of the pioneers, Kitcher describes this phenomenon of utilising different motivations and personalities as the division of cognitive labour [4]. Here, he questioned whether having a discrepancy between individual and collective rationality is beneficial. In other words, as individuals, do we always want what the collective wants? An example of collective rationality may be the exploration of *all* possible solutions for a scientific problem. In contrast, the individual may only want to find the *best* solution for the specific research problem instead of potentially examining an inferior solution. Here we see a discrepancy. Just like Kuhn, Kitcher describes and examines through his mathematical model the tension between individual and group rationality [5].

Kitcher proposes a thought experiment where scientists research the Very Important Molecule (VIM). Scientists wanted to know the structures of VIM, and two methods were available. Method one was X-ray crystallography, where they inspected the resultant photographs to examine the structures. Method two was guesswork and the building of tinker-toy models [4]. One could guess which method was favoured. All the researchers opted for method one. However, the structure of VIM was never found. Had there been somebody dividing all the scientific work, it may would have been best that that person appointed all the scientists to distinct methods. It could be beneficial to have scientists explore all methods; at first sight, some will seem inferior, and some will seem very fruitful. However, one can only be sure after exploring. In Kitcher's example, there is a discrepancy between the rationality of the individual and the community.

Moreover, history has taught us that scientific work can be erroneous. Several factors can influence the outcomes of research. A few examples of these factors can be that experimenters (unconsciously) influence subjects during the experiment. Furthermore, while collecting data, scientists can have bad luck with the sample or have faulty measurement devices [2]. Consequently, errors in research can result in outcomes that cannot be reproduced. In a Nature survey, it emerged that more than 70% of the researchers that completed the survey have failed to reproduce another scientist's study, and more than half have failed to replicate their *own* study. As a result, 52% of those surveyed agreed that a significant replication crisis is going on[6]. Factors contributing to irreproducible research were: selective reporting, pressure to publish and low statistical power. Irreproducible results may lead to false leads, and scientists might pursue less fruitful research. Therefore, replication within science is an essential part of epistemic gain.

An exciting tool for investigating these matters is agent-based modelling, which we will expand on in the following section. With the help of agent-based modelling, we will examine the division of cognitive labour in scientific research. More precisely, we will examine the ideal distribution and contributions of the replicator agent in a three-agent framework on an uncertain epistemic landscape.

The landscape in our model represents a generalised scientific research topic. The landscape is modelled as a grid of 10201 patches. Each patch represents an approach to tackle the scientific research question at hand. Some of these approaches are helpful, and others are not. On this landscape, agents can find significant and insignificant patches. Significance can be explained as the scientific potential of the approach. Significance is assigned to every patch with a Gaussian formula. The formula creates two peaks of different heights. Every time an agent lands on a patch, it gets a noisy value based on the true significance of the patch/approach. Some agents can replicate approaches. By replicating approaches, the significance of a patch may get less noisy, and therefore the assessed significance by the agents gets closer to the true significance of the patch.

The agent types moving around the landscape will react to their surroundings and try to find the most scientifically significant research approach. The surrounding includes visited and unvisited patches and the other agents. As said, we will use three agent types—the maverick, follower and replicator. The former represents an adventurous scientist. It tries to find unexplored areas of the landscape. Once it sees that approaches have been examined, it moves into other landscape areas. Moving in another direction can have the consequence of not finding significance for a while. On the other hand, the maverick can suddenly arrive in very fruitful areas not discovered before—a high-risk, high-reward mentality. The follower is more conservative than the others. It considers earlier successful research and tries to find unvisited patches around these areas. If explored approaches are more significant than its previous work, it will examine the areas around the more significant patches. As a result, it pursues meaningful yet incremental additions to science. In short, followers stay close to visited areas; mavericks move away from them. Lastly, replicator agents, as the name says, replicate approaches. They are supposed to verify all the significance that is found. One may be more confident of an approach's success by using replication. Did they succeed with the same approach twice, or was the first success a result of error? The opposite can happen as well. An approach may give a low significance at first, but it can prove itself fruitful after replicating. Replicators try to provide clarity when there is uncertainty.

In the following section, we will discuss the importance of agent-based modelling within the philosophy of science, followed by a discussion of state-of-the-art models published in the last few years. Furthermore, we will thoroughly explain our model and its simulation runs. Afterwards, we will discuss the results and the various idealisations of our model. Followed by a discussion of evaluating formal models. We will end the thesis with concluding our finding and recommendations for future research.

Importance of Agent-Based Modelling

One of the challenges of formal modelling social dynamics of scientific research is bridging the gap between the model and the target system [7]. When we look at the models we will discuss in the following chapter, there is a clear methodology for constructing formal models. First, a problem is stated: For example, the best way of transmitting data in scientific communities. Second, a theoretical framework needs to be chosen to tackle this problem. A theoretical framework can be Rational Choice Theory, which describes the best choice once a representative group is established [8]. Furthermore, if necessary, a simulation framework is needed. Some formal models do not require a simulation. A simulation can be designed using NetLogo, as discussed earlier, using agents and landscapes. Lastly, one needs to make assumptions about the real world to make designing the model possible. For example, one may think humans are rational beings and make decisions based on beliefs. This belief system can be modelled through Bayesian Reasoning. This example is an assumption we make to clarify human behaviour in formal models. These assumption decisions need to represent a phenomenon of the real world.

In recent decades agent-based modelling has become a well-established tool in the social sciences. It provides a way to study social, economic, historical or political phenomena by examining the iterated interactions of individuals that give rise to the phenomenon [9]. Examples of applications can range from agent behaviour in the stock market, the worldly spread of a new virus during a pandemic or understanding the information flow within a community. With the ever-growing capabilities and the availability of personal computers, the use of simulations based on agent models has also been growing. More recently, philosophers of science and social epistemologists have begun to follow this trend. They are also using computer simulations to model the social structures of scientific research. Examples are Zollman, Thoma, Muldoon & Weisberg, which we will discuss in the following chapter. [2, 1, 3]. Computer simulations are valuable tools for modelling the complexity of internal dynamics of scientific endeavour [7]. Klein et al. [9] discuss several advantages of agent-based modelling over other methodologies. They can provide numerical solutions of mathematical descriptions of social systems that are not tractable by using classical means. A mathematical description can help to represent and evaluate complex systems and their dynamics. An example is Bayesian reasoning, which is used to adjust the belief of a hypothesis after more evidence or information is collected. This model is based on assumptions of human behaviour, but without numerical solutions, its appliance is not waterproof. With agent-based modelling, it is, for instance, possible to examine complex social interactions over a long period without waiting for a long time for the simulation to finish. Furthermore, by controlling random effects and input parameters, modellers have the advantage of examining phenomena with high precision. Lastly, agent-based models are often modelled with their representational tools. An example of this is the use of NetLogo [10].

A further potential challenge of these models would be connecting the model with the target system using data. One thing that can make a model helpful is that it can predict outcomes and thus be able to provide recommendations. Connecting the model with data can be done by parameterisation, goodness-of-fit testing, or by any other data-oriented means [7]. Parameterisation and goodness-of-fit testing could be made more accessible with the fast-rising increase in data storage regarding this subject. For example, in the case of social network models, one can use electronic bibliographic databases, and scientific citation indexes, such as Web of Science and Google Scholar [11]. New forms of digital data in science organisations can only be more helpful in making this next step. In the following chapter, state-of-the-art models will be discussed. The models discussed include bandit problems, social networks and epistemic landscapes.

Chapter 2

Literature Review

2.1 Dealing with Uncertainty

One of the relevant factors in our model is the factor of uncertainty. One author who has tried to capture this is Kevin Zollman [2]. Zollman describes this in the way of potential error within science. According to him, there are three broad categories of error. Some occur because of sociological factors. For example, powerful political parties, cultural biases, or scientific alliances may cloud the judgment of scientific results, preventing a superior theory from being accepted. Other errors result from scientists' misconduct; data is invented, hidden, or manipulated, and as a result, others are convinced to accept an inferior theory. Finally, errors may result from simple, good-faith mistakes on the part of scientists. Everybody could have thought to do the right thing, but some random occurrences skewed the results[2]. For the progress of science, these errors can be very dangerous. A superior alternative could be dismissed, and as a result, the community proceeds with the inferior alternative. Or, science advances get decelerated because others spend too much time on inferior approaches.

Errors are everywhere, and mistakes can occur without the scientist knowing. Sometimes these errors may lead to what seems to be very successful results. By inventing data, scientists can let the community believe their research is superior. Due to political preferences, research may not get the funding they need, and then an approach is never to be explored. Finally, due to honest mistakes, an experiment can look more or less promising than it is. A scientist that gets to know the results of an experiment may be influenced by the outcomes of that experiment. It may cause a new scientist to avoid trying this approach because it does not seem to work, or vice versa. For this reason, Zollman argues that it might be beneficial for a community to have limited information on what everybody's result is. If one does not know about an experiment, one is not influenced by it. He demonstrates this with his model about transient diversity [2]

2.1.1 Bandit Problems and Social Networks

Zollman's model has two main components: social networks and (learning in) bandit problems. Firstly, we will discuss the latter. With the bandit problems, there is the gambler and two slot machines. The gambler is a metaphor for the scientist, and the slot machines are a metaphor for scientific theories. Each time the gambler chooses a slot machine to play with, the gambler observes and remembers the outcomes. For simplicity, "win" or "lose" are the only options. The gambler should learn the win probability of each slot machine by playing. He will probably try out both slot machines initially, but when should he stop? How secure does the gambler need to be to choose only the slot machine with the highest payoff? This choice gives the dilemma of one wanting to gain information but also wanting to play the slot machine with the highest probability—also known as the dilemma of exploring versus exploiting. In this model, we assume that past success or failure does not influence the success or failure of the experiment in the future.

In Zollman's model, the scientists will be presented with two potential methods, each with an intrinsic probability of success. The scientists try to find the method with the highest intrinsic probability of success and will pursue this method. They will update their beliefs using Bayesian reasoning based on their success and the success of others [2]. The way Zollman chooses to model this type of learning is with a beta distribution. One can use a beta distribution as a prior, and after sampling and updating, one can use the beta distribution as a posterior[2]. The posterior mean is given by:

$$\frac{\alpha + s}{\alpha + \beta + n}$$

where α and β are the priors, n the amount of trials and s the amount of successes. The scientist's beliefs will eventually approach the true mean by doing enough trials.

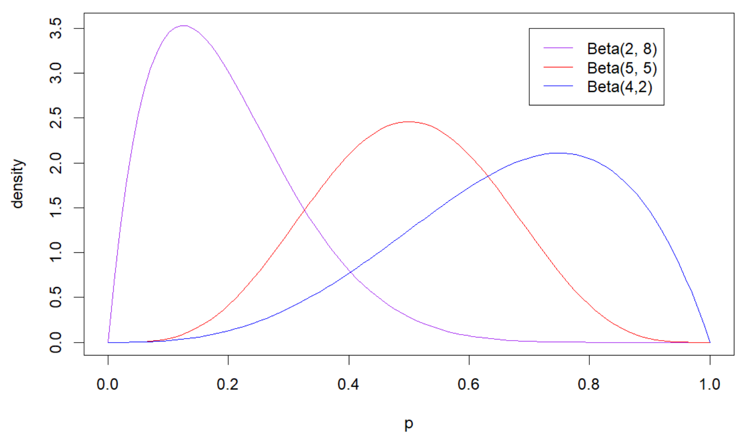


Figure 2.1: A graph showing beta distribution for different prior values of α and β. Higher values for α give higher chances of success. Higher values of β give higher chances for failure

Scientists do not only learn from their work; they can learn from the scientific community as well. Central to this are social networks. The social network is the second component used by Zollman [2]. A social network is a set of individuals represented as nodes. These nodes are connected by edges. The edges represent how information can transfer from one scientist to another. Zollman experimented with three social networks, the cycle, wheel and complete graph. The different forms of social networks can be seen in fig. 2.2. In the complete graph, everybody can communicate with everybody. The wheel has one scientist in the middle that can share information with everybody, but the others can only communicate with their neighbours. In the cycle, scientists can only communicate with their two neighbours, left and right from them.

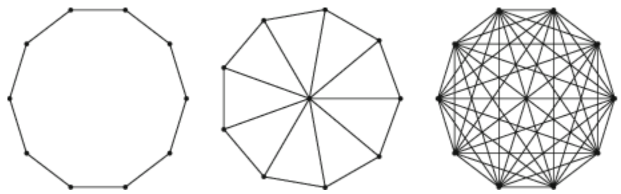


Figure 2.2: From left to right: A cycle, wheel and complete graph

Results of the Model

Zollman experimented with two different factors. The first was the limitation of information by having more or fewer edges between the nodes, and second, he experimented with varying the priors in the beta distribution. The former will be discussed first. The three networks, as discussed earlier, were tested, the cycle, wheel and complete graph. Here, Zollman found a surprising result. Information appears to be harmful in terms of identifying the superior theory. [2]. The cycle is superior to the wheel and the complete graph, which means that the probability of successful learning is the highest with the cycle.

The other part of the experiment was altering the priors. The priors determine the strength of one's individual beliefs. With higher priors, there is less variance. The means of the beta distribution can remain the same, but the chance of drawing other values is smaller. This has as a result that scientists with high initial priors are more resistant to initial evidence. Zollman found that the initial superior networks reversed as the initial parameters increased. The complete graph was far superior in successful learning than the cycle and the wheel. With the low initial parameters, the results stayed the same. This might be because when the initial parameters are up, much new information is needed to change that belief. When initial parameters are down, an agent is easily convinced by only a few outcomes of others.

Uptake

What can be concluded from these results is that scientists either need not be fully informed by all the other scientists, or their prior beliefs should be high enough to secure that scientists will not discard their potentially superior scientific method too fast. Although, it is essential to state that both should not co-occur. High prior beliefs and uninformed networks were inferior to the others as we recall the previously discussed results. As is the case, vice versa. Arguably, it is less likely one can influence an scientist's priors, so it would make sense to create a structure of science that is less informative.

2.2 Division of Cognitive Labour

2.2.1 Followers and Mavericks on a Epistemic Landscape

Michael Weisberg and Ryan Muldoon simulate the division of cognitive labour with a computational agent-based model. They designed and tested their model in the NetLogo environment. In their paper, they answer what the optimal distributions of cognitive labour are [1]. For this, Weisberg and Muldoon adopted an agent-based approach. These agents represented different scientists with contrasting attitudes. Agents are able to move on what they call the epistemic landscape [1].

Epistemic Landscape

The landscape itself is a broadly construed research topic. The different positions on the landscape represent the slight differences in approaching this research topic. We will later discuss how these approaches can differ and what this means regarding positions on the landscape. Furthermore, the epistemic landscape consists of several components. We will discuss the components in the following.

As we have established, the epistemic landscape corresponds to a specific research topic that engages a group of scientists [1]. Topics one can think of are thermodynamics in physics [12], plant chemical communication in biology [13], or image processing in the brain in neuroscience [14]. The simulation landscape represents one research topic. The second component of the epistemic landscape is the approaches. The narrow specifications of scientists of how a single scientist or scientific group investigates the research topic. An approach can consist of [1]:

- The research question being investigated
- The instruments and techniques used to gather the data
- The methods used to analyse the data
- The background theories used to interpret the data

For example, among the researchers studying the causes of peptic ulcer disease (PUD), there were scientists with different instruments and techniques. To detect particles in chemistry, chemists often use a specific stain. There are different kinds of stains. Some used a silver stain, and others a gram stain. Researchers in this example had reasonable grounds to choose these stain types. However, different stains will have different outcomes as a result. Both researchers used the same data-analysing methods and the same research question. Nevertheless, the different stains will give distinct images. Therefore, their outcomes will be different. The case of PUD will later be discussed in a more detailed manner. We use this example as an introduction to our model. Most importantly, every approach component will influence an experiment's outcome. In the model, this amounts to different locations on the landscape. These locations can be expressed through x and y-coordinates.

A last component Muldoon & Weisberg describe is epistemic significance. Most of the patches on the landscape have significance. A Gaussian formula assigns the significance values to every patch. The formula creates two hills with peaks. Because the values are assigned through a Gaussian distribution, there are two hills with a peak and decreasing values surrounding the peak. the foot of the hill has significance values starting from one.

There are dividing views in classical philosophy of what exactly is scientific significance. Some argue that true facts can have intrinsic value, as opposed to this is the view that the judgment of scientific significance is merely the result of dominant ideologies and other social forces that influence scientists in their research. Weisberg and Muldoon did not commit to one specific view. Most important is that all agents in their simulation try to find the highest value on the landscape. In other words, find the top of the highest hill.

From these three components, the authors construct the epistemic landscape. The scientific topic is represented as the whole landscape/grid. An approach is one patch located on the grid, and some of these patches contain scientific significance. To find the highest significance, some distribution of agent types is needed. In the following, we will discuss the two distinctive agents. These agents are important for this thesis, the follower and maverick agents.

Followers and Mavericks

As stated, Muldoon & Weisberg have created an agent-based model. Moreover, the goal of this model is that agents on the grid learn from the results of themselves and other agents. Learning about the surrounding will influence the behaviour of the agent, which will influence the behaviour of the other agents. How they learn is described in the rules of the specific agents. This model's two important agent types are the follower and maverick agents. The former behaves as a conservative scientist that learns from research related to their research or incrementally improves their own research. The conservative scientist, intuitively, is how science typically progresses. Scientists learn about other studies or have a successful study by their selves and try to expand on this with minor changes. This way, the epistemic realm gets expanded step by step. The latter is a more experimental scientist who tries to deviate from the work done before him. To be a maverick, one has to take risks. When a maverick sees that an area is already explored, it finds another area that has yet to be investigated. A change of direction can lead to a scientific breakthrough, or it might lead to nothing. Either way, this behaviour might help other agents. It can show paths that need further exploring or show that an area is insignificant.

Formally, follower agents observe their neighbourhood. An agent's neighbourhood are the eight patches surrounding the current patch of an agent. First, the followers check whether the surrounding patches are visited or not. If so, they check if the significance of that patch is higher or lower than their current patch. It does so for all the surrounding patches. If there are one or more patches with higher significance, it goes to the highest one. Otherwise, they will go to a patch that is not yet visited and in their neighbourhood. This will continue until the agent finds the highest significance or when no patches are available.

Mavericks, on the other hand, avoid previously examined approaches, while followers emulate them [1]. Mavericks examine whether their current significance is higher or equal to their previous significance. If this is the case, they look for unvisited patches. If there is an unvisited patch, mavericks move in this direction. They will take the patch with the highest significance if no unvisited patches are left. When the current patch has a lower significance than the previous significance, mavericks move one patch back and take another random direction. These rules represent the attitudes of conservative and experimental scientists.

Because mavericks can keep walking on the grid, they will help other follower agents find paths to the significance peak. The follower will follow the same path as the maverick that found significance, and eventually, it has a higher chance of finding the highest significance.

Results & Uptake

Muldoon & Weisberg's measure of success is that of epistemic progress. Epistemic progress is defined as the proportion of patches with significance that is found against all the patches with significance. Muldoon & Weisberg found that adding only one maverick significantly helped epistemic progress. After this, they tested different ratios of mavericks to followers, finding that a 50/50 division is most beneficial.

This shows that mixed strategies of different agents are likely to be helpful in terms of finding the highest significance as a community. This results from mavericks finding new paths for the followers in a daring matter, making it easier for followers to travel across the grid.

2.2.2 An Extension of the Epistemic Landscape

Johanna Thoma criticises the assumptions and outcomes of the model. She discusses these wrong assumptions in her work and tries to improve the model by expanding and correcting with her own assumptions[3]. Thoma designed and tested her model in the NetLogo environment as well [10]. She revisits the division of labour and plays with the notion of flexibility and ignorance. Thoma argues that Muldoon & Weisberg cannot show that it is beneficial to have mixed strategies. When looking at their results, it can be seen that the homogeneous group of mavericks is superior to the mixed groups and homogeneous group of followers due to the movement strategies of the agent rules described by Muldoon & Weisberg. Full groups of mavericks in Muldoon & Weisberg's model are faster and more complete in finding significant patches. Before we look at Thoma's extensions and changes, similarities between the models will be discussed.

Similarities of the Models

To start with the landscape. As Muldoon & Weisberg, Thoma's model is a grid with x- and y-coordinates. Every (x,y)-combination represents an approach, and all approaches consider a specific scientific research topic, as discussed before. Apart from a coordinate all patches have a significance value. The landscape has two significant peaks, the highest and most optimal, and another sub-optimal peak.

There are still two different kinds of agents used in Thoma's model. Muldoon & Weisberg called these the follower and the maverick agents, and Thoma called these the extractor and explorer agents, respectively. For the sake of clarity and readability, the name types of Muldoon & Weisberg are being kept. The agents' attitudes are supposed to be similar, the follower, being the more conservative scientist, is trying to expand its field to find helpful research and add small adjustments to expand the epistemic realm. And the maverick, the agent that follows the motto: high risk, high reward. The maverick tries to avoid explored areas because he or she wants to find something new. Both models also represent these attitudes, though Thoma has some considerations about the Muldoon & Weisberg agents, which we will discuss in the following section.

Changes to the Model

Firstly, the follower agents in the Muldoon & Weisberg model seem to get stuck on sub-optimal positions and go back and forth till infinity. This behaviour can be seen in fig. 2.3.

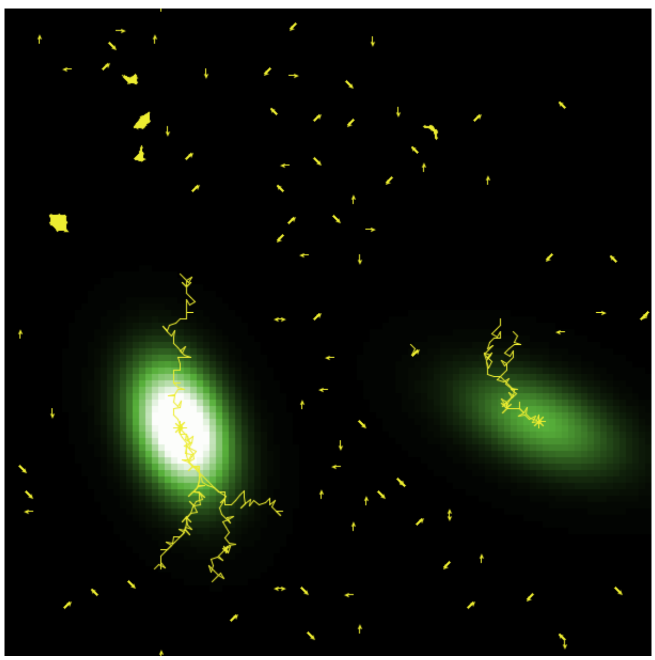


Figure 2.3: A simulation coming from the Muldoon & Weisberg model [1], indicating agents getting stuck on sub-optimal patches. What can be seen here is that some agents (in yellow) stay on the same position and making a yellow dot. Other do find the significant (white) patches.

For Thoma, this is not rational. When scientists do research that does not work out as well as they hoped, they will try another experiment and see whether this new experiment is more significant than the approach before. If not, they would never return to their previous inferior work because this would be costly and ineffective. Therefore, Thoma argues that her follower agents should never redo research, hence never get stuck in their current positions. In the model, this amounts to followers being able to keep moving. When Thoma's follower finds a visited patch in its neighbourhood with higher significance than the current patch, it moves to the closest unvisited neighbouring patch. This way, the agent does science that is likely to be successful but still slightly different than has been done before. The maverick, however, behaves like a hill climber. Unless it gets on a less significant patch than the previous one, it will move straight ahead. When it would land on a patch that has been visited before, the explorer moves in another direction. This is because the explorer wants to have significant discoveries that are dissimilar to the work that is done before.

A second characteristic that Thoma changed is that of local and non-local movements. Thoma argues that taking just one step at a time, moving one patch at a time, represents inflexible and ignorant scientists. For Thoma, it is not likely that scientists can only *slightly adjust* their research by adjusting a small amount of the equipment, altering the research question, or having only minor differentiating theories. Furthermore, Thoma argues that scientists are reasonably knowledgeable about other researchers' work, not only in their direct surroundings but also about what is slightly out of their area. This makes the scientist more flexible and less ignorant about others' work. Consequently, both agent types should be able to make bigger steps; in Thoma's model, agents have a movement range between one and ten patches.

Results

As a measure of success, Thoma recorded what proportion of the total significance of the landscape each population discovered after various periods [3]. Thoma also recorded how much significance each type of scientist discovers. This measure shows how productive each type of scientist is. This is slightly different from Muldoon & Weisberg. They tested the proportion of significant patches that was found. They did not pay attention on how significant a patch was.

There were three categories to test the local and non-local movement: local, medium range and global movement. During local movement, all groups with a minimum size of 40 found approximately the entire epistemic significance of the research field within 500 rounds. However, scientists do not only care if they find all the significance; it would be better if they could find it more quickly. How bigger the group was, the faster the group was to find all the significance. Furthermore, homogeneous groups of mavericks do better than homogeneous groups of followers and mixed groups with the local movement.

At the medium-range movement from a range of 3 upward mixed groups are starting to do best, followed by the homogeneous groups of followers and mavericks doing the worst. Also, larger ranges of movement are beneficial for finding the most significant. In addition, larger ranges of movement mixed groups are even more superior to the other groups than before [3]. This suggests that the division of labour between followers and mavericks is beneficial when the movement is not local.

To examine global movement, Thoma had to alter the maverick rule slightly. Mavericks randomly picked an unvisited patch on the landscape. Diversity in labour is still beneficial with global movement, but is slower in finding the entire significance. On average, followers are more productive than mavericks. At the range of ten for movement, followers were four times more productive than mavericks.

Uptake

As opposed to the Muldoon & Weisberg model, Thoma's model shows that it is beneficial to have mixed strategies when movement is not local. When scientists are uninformed and inflexible, in other words, when their movement radius is too small, follower types are less likely to follow the maverick types to more fruitful areas of research [3]. On the other hand, when scientists are too free, some fruitful areas may be ignored. A second finding shows that follower agents are more productive in finding significance than maverick agents, in contrast to the Muldoon & Weisberg model, where it seems that it is always better to have a whole group of mavericks. Thoma's model seems better defended, both for having more realistic assumptions and for model behaviour and results. Only some people can be mavericks because the risk of getting no reward is too high. However, it is beneficial to have mavericks in the scientific community. Thoma argues that special incentives are needed to make scientists behave like maverick agents.

Chapter 3

The Model

3.1 Peptic Ulcer Disease

The Peptic Ulcer Disease research case is often used as an example of scientific inquiry gone wrong [15, 16, 2]. Others have used this case to illustrate the information flow within social networks. We will use it as an example for the division of cognitive labour. It shows that it would have been beneficial to have scientists replicating experiments to recognise the superiority of scientific methods.

Peptic Ulcer disease (PUD) causes ulcers in both gastric and duodenal (beginning of the intestines) areas. The predominant symptom of PUD is gastric pain [17]. Among others, Zollman [2] described two leading hypotheses for the cause of PUD; hyperacidity (a medical condition where the stomach creates too much acid fluids) or a bacterial cause. However, Zollman stated that the bacterial hypothesis was abandoned for nearly 30 years because the hypothesis was thought unfruitful by the scientific community. The abandonment has been explained by the influence of Palmer's large-scale study in 1954 [18]. His research showed that no bacteria could be found in the gastric areas, disproving the bacterial hypothesis. Palmer's finding was later proven to be false by Warren and Marshall in 1983, telling us that a bacteria indeed caused the Peptic Ulcers [19]. It was a specific type of bacteria strand—Helicobacter Pylori. The acid covering the ulcers can cause pain. The combination may be why acid was thought to be the leading cause of peptic ulcers.

According to Radomski [15], the story has two flaws. Firstly, it was falsely believed that either bacteria or acid caused PUD. Secondly, the impact that Palmer had on the subject was overestimated. [18]. Before, it was interpreted that Palmer had a massive contribution to the matter. Initially, it was thought Palmer influenced the other scientists so much that nobody thought it worth pursuing the bacterial hypothesis. Therefore, many argued that others had abandoned a superior theory because of his influence. For this reason, Zollman argued that if there had been less information flow, others would have pursued their research and then had a chance to find the actual cause of PUD a lot earlier [2].

However, Radomski argues that the bacterial theory had been abandoned sometime before Palmer's large-scale study [15]. Palmer's research was far less influential and may have been more like the last nail that closed the coffin. What Radomski has shown is that the problem at hand had been much more fine-grained. At that time, different theories existed about the causes of PUD. These included: Hyperacidity, the excess of acidity in gastric and duodenum areas. Second, stress-induced PUD; for example, the increasing numbers of PUD cases during WWII, were used. The increases were linked to increased stress levels during the war. Third, PUD-types, it was thought certain people were inclined to produce these ulcers more than others. Finally, bacteria-induced ulcers. However, the bacteria hypothesis was abandoned around ten years before Palmer's research.

A plausible reason for this abandonment could be that although scientists had some proof for certain types of bacteria causing the ulcers, others could not reproduce the results of these experiments. One exception was the research of Freedberg and Barron [20]. They found the specific strand of bacteria causing PUD, Heliobacter Pylori. However, their research was small and therefore overlooked by others. Interestingly, Freedberg and Barron advised against using a specific stain to detect bacteria in gastric environments. In biology, dyes are used to stain cells. Dyes bind with specific parts of the cells, which increases contrast. The increased contrast helps with the visibility and distinguishing of different cellular parts [21]. Barron and Freedberg advised everybody to use a silver stain instead of the gram staining technique; both stains create different colours and bind to different parts of the cells. Making bacteria invisible for one and visible for the other. However, the gram stain was used by Palmer, even though he cited Freedberg and Barron in his large-scale study. For some reason, this was ignored/unnoticed by Palmer.

What can we learn from this specific case? More importantly, how is it applicable to our model? We have seen that multiple approaches and theories were used to examine the causes of PUD. Some were experimenting with different strands of bacteria. Others were examining the cause of acidity on peptic ulcers. Another group tested the influences of stress, and the last one claimed there was a specific PUD type of people, the people that were more likely to get peptic ulcers.

A few of these approaches were similar, but most of them were very different from each other. For instance, one may try to prove the bacteria hypothesis but looks at the wrong bacteria strand. Trying to find another strand of bacteria as the cause is almost similar but a slightly different approach. A wholly different approach is trying to prove it is induced by stress. The fine-grained distinct approaches help us build an epistemic landscape as a grid. Namely, a landscape with a research problem (finding the cause of PUD) with all the possible approaches to tackling the problem. Some of the approaches were judged as significant. In this case, the acidity hypothesis, and some were less significant—for example, the research of Freedberg and Barron. Although Freedberg and Barron were later right about the type of bacteria, their research was found too small and inconclusive [15]. Therefore, it was thought not worth pursuing this approach.

In this case, it would have been beneficial to have some open-minded replicators trying to reproduce the findings of Freedberg and Barron earlier. It was necessary to verify how good this approach was. For instance, by increasing the scale of the experiment. Freedberg & Barron found a solution to the research problem. Nevertheless, their research was small and overlooked. Others did not see the significance of this study. Therefore, their approach was not pursued, and followers moved away from this area because other approaches were thought better than Freedberg & Barron's research. Had there been a replicator, increasing the study's sample size and trying to duplicate the study's results, then there was the chance that this replicator found more significance in this approach. This reassurance would make it more tempting for others to pursue.

Also, it might have been beneficial to have mavericks explore the areas in times of abandonment. Revisiting abandoned areas is essentially what Warren en Marshall did. Sometimes, mavericks do not find significance at all. When this is the case, they are potentially still helpful for the scientific community. They show that some areas are insignificant and that other agents do not have to follow these paths. In the case of PUD, too many theories were abandoned too early without replicating to find the true significance of the approach. Had there been somebody that replicated the Freedberg & Barron study a lot earlier, we may have found a solution for PUD earlier.

In the following, we will describe the components of our model. Here, we will discuss the behaviour rules, the motivation for these rules, the performance measures and the setup of the simulations.

3.2 Agents in an Uncertain Environment

NetLogo was used to design and simulate the agents and landscape. [10]. NetLogo makes it possible to design and visualise multi-agent models. Therefore, researchers and academics use it to simulate complex multi-agent phenomena. This work is partially built on the work of Muldoon & Weisberg [1]. Most components are adjusted and/or expanded, except for assigning significance to every patch; this model uses the Gaussian formula as they did.

In the previous, we saw two takes on epistemic landscape models. Muldoon & Weisberg, and Thoma examined the ideal distribution of labour with two different agents. Namely, the distribution between followers and mavericks. These agents have different attitudes. The followers are more conservative than the mavericks; they want to stay close to what they and others have done before to expand the area and their knowledge incrementally. On the other hand, mavericks are more adventurous than followers. They move to areas of the landscape that are not explored before.

In Muldoon & Weisberg's model, agents could only take one step at a time. It was also possible for agents to go to the same patch they had already visited. However, revisiting the same patches resulted in followers getting stuck at insignificant patches. In Thoma's model, agents were not allowed to go to the same patch, as she argues that it is illogical to redo unsuccessful research. She added that agents were also allowed to take bigger steps within one round. Taking more steps at a time represents that agents are more flexible and knowledgeable about their surroundings. According to Thoma, Muldoon & Weisberg found that homogeneous groups of mavericks are superior to any other mixed or homogeneous groups of agents. After Thoma changed the model, she found that mixed groups of mavericks and followers are superior to other groups, when agents have a non-local or global range of movement. Thoma found that agents need to have a minimal range of three to be superior for mixed groups to be superior.

Both models assume that when agents arrive on a patch, they immediately know their approach's significance. However, how can they be sure? In Zollman [2], we found that research can be full of errors. One can, for instance, use the wrong equipment for the experiment, the equipment can be faulty, or one could have bad luck picking a sample of the population. Measurement error adds noise to predictions, increases uncertainty in parameter estimates, and makes it more difficult to discover new phenomena or to distinguish among competing theories [22]. It is often believed that when one achieves statistical significance with noisy conditions, the observed effect would only be more significant in perfect conditions [22]. But this cannot be assumed. In the previously discussed Nature survey, we saw that more than 70% of the researchers who tried to replicate a study failed to reproduce the outcomes. Failing to replicate even happened for more than 50% of the survey participants for their own research. Failing to replicate does not always have to mean that the outcomes of a study are false. Nevertheless, one should be cautious when it happens.

Therefore, we argue that science is less certain than Muldoon, Weisberg, and Thoma describe in their model. Immediately finding the true significance is too much of an idealisation. Consequently, uncertainty is added to our model to make the model more complex. Because of the addition of uncertainty, there needs to be something to make the uncertainty less noisy. We will do this with the arrival of a new agent: the replicator. In the following, we will discuss what the epistemic landscape of our model looks like, what changes we made to the already existing agents and what the behaviour is like of the new agent, the replicator.

3.2.1 The Uncertain Epistemic Landscape

As Muldoon & Weisberg, a 101 x 101 landscape is used. This results in 10201 distinct patches. These patches represent different approaches. Neighbouring patches are very similar in approach; patches on opposing sides of the landscape are very different. Then, every patch is assigned a significance score. It is assumed that significance is not distributed randomly on the landscape but in two Gaussian-shaped hills with single peaks. One peak is higher than the other. The Gaussian formula randomly selects a random x- and y-coordinate as the centre point of the hill and creates the rest of the hill around this point. To create a second hill, it does the same process again. Most of the fields will have a significance score of zero. The highest significance can vary from around 500 to 2000. Significance represents the scientific potential of the patch. Hence, the agents want to find the highest significance on the grid.

Uncertainty Layer

As discussed, errors can occur unknowingly in scientific research. Because of this, agents cannot immediately determine an approach's significance. Instead, whenever an agent lands on a patch, it is like the agent pulls a lever from a slot machine, and the slot machine will give the agent a number scattered around the true significance. To draw a number, one can make use of different distributions. We discuss two of them; the beta-and normal distribution.

With Zollman, we discussed how to deal with uncertainty. Zollman used a beta distribution to come closer to the true mean. The beta distribution has the priors alpha and beta. Both can be set to one or any other number. These priors represent how certain one is about their methodology for an experiment. Low priors can be seen as very unsure, and high priors are sure about their research. These priors are subjective and not necessarily linked to the truth. In other words, it may have a factual basis but can also mirror the subjective strength of belief. After the priors are set, one can count the successes and the failures of an experiment and use the formula:

$$\frac{\alpha + s}{\alpha + \beta + n}$$

This will give the posterior mean where s are the successes and n the number of the same experiment executed. The beta distribution has proven to approach the true probability efficiently. Although this helped Zollman's model, where he only experimented with whether one methodology was superior to the other, this distribution may not be helpful for our model. The Beta-distribution has one major problem concerning our research. The beta distribution focuses on binary outcomes of one specific research. An experiment is successful, or it is not. Alternatively, Zollman's case focused only on two rivalling methods. One method was superior to the other, and after re-examining both methods many times, one better understood which method was superior. Our model has 10201 distinct approaches, and it is much more fine-grained regarding the approaches to tackling a research problem.

Consequently, another distribution is used. Namely, the normal distribution. The normal distribution describes a family of continuous probability distributions. These distributions have the same general shape and differ in their location. Their mean determines the location on the x-axis, and the standard deviation determines their scale parameters. The graph has a symmetric, and bell-shaped curve [23]. The formula for the normal distribution is given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

with μ being the mean and σ the standard deviation. When there is a mean of 500 and a standard deviation of 150. The chances of drawing the number 500 are much higher than the number 150. However, chances of drawing 150 are higher when the standard deviation increases. This can be seen in fig. 3.1.

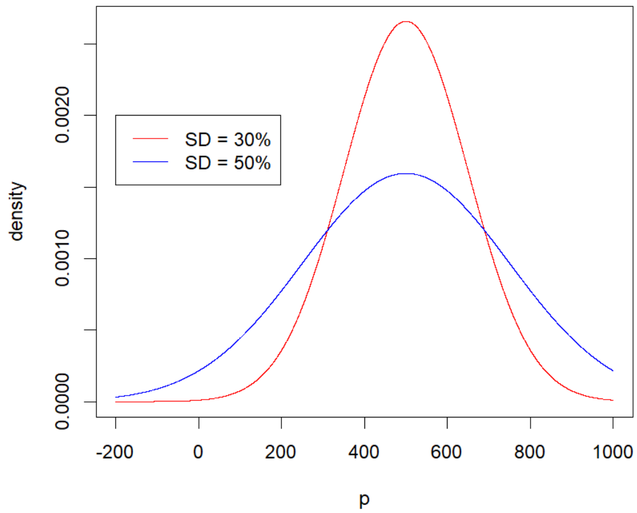


Figure 3.1: Normal distribution for means of 500 and standard deviations of 30 and 50% of the means. One can see that the variance is bigger with a higher standard deviation

In this model, we first assign the true significance to each patch. The patch with the highest significance can be between 500 and 2000 significance. Around the top of the peak, the significance will decrease until the bottom of the hill. Many patches on the grid will have a value of zero. Then, every time an agent lands on the patch, it will draw a number from the normal distribution. The distribution has the true significance as mean and 30% or 50% of the true significance as standard deviation. We have chosen a for a percentage of the true mean because we think that the potential deviation should be proportional to the true value. This means the higher the true significance, the bigger the potential variance, which is also demonstrated in fig. 3.1.

Subsequently, agents will be sprouted randomly over the grid and land on different patches. Every time an agent lands on a patch, the lever of that gets drawn again. If the lever of one patch has been drawn several times, the agents will assess a patch over the average of all draws.

The uncertainty may influence the behaviour of the agents. One can find a high significance first, which is actually low and will stay in the surrounding area. Or vice versa. Because of this, agents may move in the wrong direction. The purpose of the replicators is to avoid this. By replicating, everybody should have a good understanding of what approach is significant and what approach is insignificant.

In the following, the different agents will be discussed. The follower, maverick and replicator. What is similar to the agents of Muldoon, Weisberg and Thoma? What adjustments and what is added to the model regarding agent types and agent behaviour rules? Furthermore, we will discuss the landscape changes' consequences on the agents.

3.2.2 Three Different Agent-Types

To find the most significant approach, searching agents are needed. Like Muldoon & Weisberg, this model uses the follower and maverick agent, and like Thoma, this model extends their behaviour with non-local movement. To do this, the behaviour rules of the agents from the Muldoon & Weisberg model needed to be changed. What is added is a new type of agent. This agent replicates approaches and is referred to as the replicator agent. In short, replicators need to reduce the noise of an approach's significance; they can move to approaches that have been tried before and replicate this specific approach. The following will discuss the behavioural rules for the agents and how they work in combination with the uncertain landscape. Firstly, the follower agent.

Followers

As discussed before, followers would like to be in the vicinity of already successfully done research. It looks around, sees the best approach known so far and does something similar. With this, they aim to find better approaches gradually. For researchers, but also in day-to-day life, this is a very reasonable thing to do. We find things that work, reason about it and try to improve them step by step. That is the attitude of the follower agent. Shortly, in Muldoon & Weisberg, followers look in their (Moore) neighbourhood; if nothing is visited, their followers go to a random patch. If patches are visited, they look at whether one or more have higher significance than the patch they currently are. If that is the case, their followers move to the patch with the highest significance [1]. This results in a lot of the follower agents that take the same path to the eventual goal. Followers are precisely following other agents. Consequently, many agents get stuck, as in fig. 2.3, and much of the landscape is not explored. Therefore, their behaviour was adjusted not to visit the same patch as others. Like Thoma [3], followers do not revisit the patches already visited before. Thoma argued that it is illogical to revisit a very bad approach. Revisiting bad approaches is what happens in the Muldoon & Weisberg model. In our model, revisiting explored approaches is the main task for the replicators. It will add noise when followers can revisit patches as well. In exceptional cases, a follower returns to the previous patch to go in another direction. Returning to patches only happens when no options are left in the current position.

The other aspect is the moving and perception range of the follower agents. Thoma argued that by letting agents only take one step at a time, agents are inflexible and ignorant about their surroundings [3]. We can assume these approaches are very fine-grained if we look at the landscape of 10201 distinct patches representing a distinct approach. Meaning a patch adjourning another patch is a very similar approach to that of the current patch. It is not unreasonable to think that researchers should be able to perceive a bigger range of patches and notice that somewhere else (I.e. four patches further), a successful approach has been applied to the current scientific problem. Just like Thoma's agents, our agents can look at patches in a chosen range. We can alter this range from one till ten patches. Regarding the neighbourhood and range, we considered two different ways to look at the range of the neighbourhood. Those are the von Neumann neighbourhood and the Moore neighbourhood. Both will be discussed because both will have their purpose in our model.

The Moore neighbourhood is a square-shaped neighbourhood that defines a set of cells. The surface of the neighbourhood can be determined by:

$$N^M_{(x_0,y_0)} = \{(x,y) : |x - x_0| \leq r, |y - y_0| \leq r\}$$

In fig. 3.2, one can see a Moore range with different values for r. The number of cells in a Moore neighbourhood is $(2r + 1)^2$.

The von Neumann neighbourhood is diamond-shaped. We can define the surface with the following formula:

$$N^v_{(x_0,y_0)} = \{(x,y) : |x - x_0| + |y - y_0| \leq r\}$$

Next to the figure of the Moore neighbourhood, we can see how the von Neumann neighbourhood is shaped. We can calculate the number of patches by $2r(r + 1) + 1$, With r as the radius. The Moore neighbourhood gains more patches with an increasing range than the von Neumann neighbourhood. With an r of 2, the von Neumann neighbourhood gives a surface of 13 patches, while the Moore neighbourhood has 25 patches. With a range of 5, the von Neumann neighbourhood has 61 patches and the Moore neighbourhood with 121. Alternatively said: the Moore neighbourhood is almost twice as big compared to the von Neumann neighbourhood. Whether we use the von Neumann or Moore neighbourhood greatly impacts the range of our follower agents. Horizontally and vertically, they can reach the same patch, but diagonally the Moore neighbourhood can reach much further.

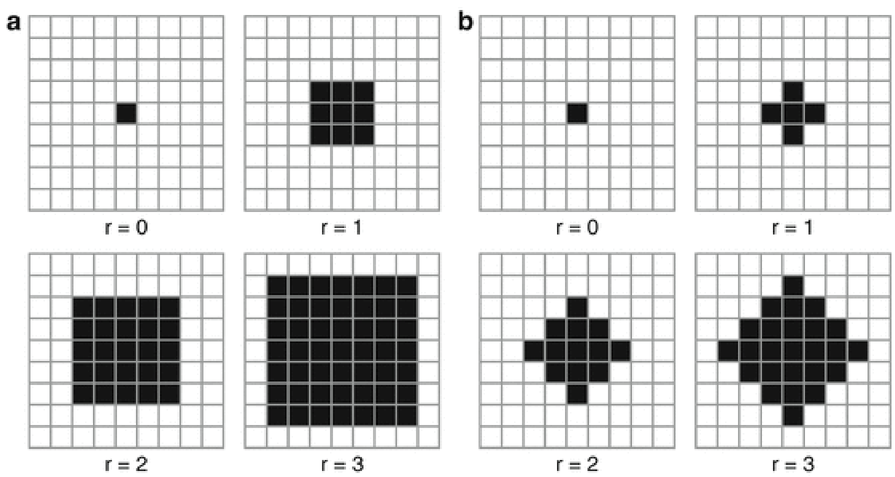


Figure 3.2: On the left we see the Moore-neighbourhood and on the right the von Neumann neighbourhood. One can see that the surface of the Moore-neighbourhood is a lot bigger while having the same range.

For the followers, we chose the von Neumann neighbourhood. The landscape represents approaches. For example, when the agent goes one step north, this can be seen as a change in equipment, while going west can be an alternative way to analyse the data obtained. So, when the agent goes to north-west, it changes its equipment and its analysing method. For the von Neumann neighbourhood, this is seen as taking two steps. If the follower has only a range of two, the follower needs to stop moving. For the Moore neighbourhood, this means that the agent took only one step. With a range of two, followers can travel twice as far in a Moore neighbourhood. The range of the neighbourhoods can be set before every run. It can be set from one to ten. The range stays the same during the run. The following is a more detailed description of the follower rules:

- **Ask:** Are there patches with an significance in the current Von Neumann neighbourhood?
  - **If no:** Are there any unvisited patches in the von Neumann neighbourhood?
    - **If yes:** Go to one of the unvisited patches.
    - **if no:** Stay on the same patch.
  - **If yes:** Does this patch have a higher average significance than my current patch and is this the highest average significance that can be found in the neighbourhood?
    - **If yes:** Does this patch have a neighbour that is unvisited yet?
      - **If yes:** move to one of the neighbouring unvisited patches
      - **If no:** Move to the a unvisited patch at a minimal distance from the agents current position.
    - **If no:** Find a patch that is unvisited and move there. Otherwise stop searching

As one can see, the follower looks for more significant approaches than its current approach. If there is no such patch available, the follower will carefully expand their region to find higher significance by their selves or higher significance that is found by others. Because the Gaussian formula creates two peaks, we know that significance is found close to each other.

We assume that this is also applicable to normal science. One will find only slightly different results if one uses a slightly different method. Therefore, we have decided for the followers that it is reasonable to visit a neighbour of the patch with the highest significance. However, all the neighbours may be visited already. We want followers to keep doing research. To avoid stopping on a patch, followers are allowed to find and go to the unvisited patch closest to their current positions. This way, they will remain in areas where significance is likely to be found. When they reach the closest unvisited patch, they are forced to explore the area again and find highly significant patches. When followers discover that the highest significance around them is lower than the current patch, it may indicate that heading that way would mean they are going in the wrong direction with the remaining area unexplored. Therefore, the followers have the rule to find another unvisited patch in a random direction to explore their surroundings more. It is plausible that an area is fully visited. When all unvisited patches are out of reach, the follower will stop looking and stay on the same patch.

At last, the followers do not know the true significance of a patch. They will get their score by drawing a number from the normal distribution. A consequence of this can be that the highest significanct patch in the von Neumann neighbourhood may be perceived as higher, but in reality is lower. This would result in that specific follower going in the wrong direction. As a result, it may take longer to get to the patch with the highest significance. Now that we described the rules of the follower agent, we will discuss the maverick agent.

Mavericks

The maverick agent tries to find new approaches in unexplored areas on the epistemic landscape. It does not do so while staying in the same part of the landscape. However, it does so by going straight ahead as long as it finds significance equal to or higher than the previous patch. It operates as a hill climber. Furthermore, when the maverick cannot go straight ahead because all in front of it is visited, it will choose another direction as the maverick wants to explore other regions than other agents. Although it may not be the primary goal of the maverick, these rules may pave the way for other agents, especially for the followers. When followers are near maverick paths, they can now observe whether one of the visited patches is of higher significance than the follower's current patches. It could also mean that mavericks will pursue approaches that are low in significance for a longer time. Nevertheless, they would still be helping the scientific community, as discussed by Kitcher [4]. If mavericks do not find any significance, they may help by showing what is not significant on the landscape. Muldoon & Weisberg said: *Like followers, mavericks take in to account which approaches have been previously explored and which ones were successful. However, unlike followers, mavericks avoid previously examined approaches, while followers emulate them.*[1].

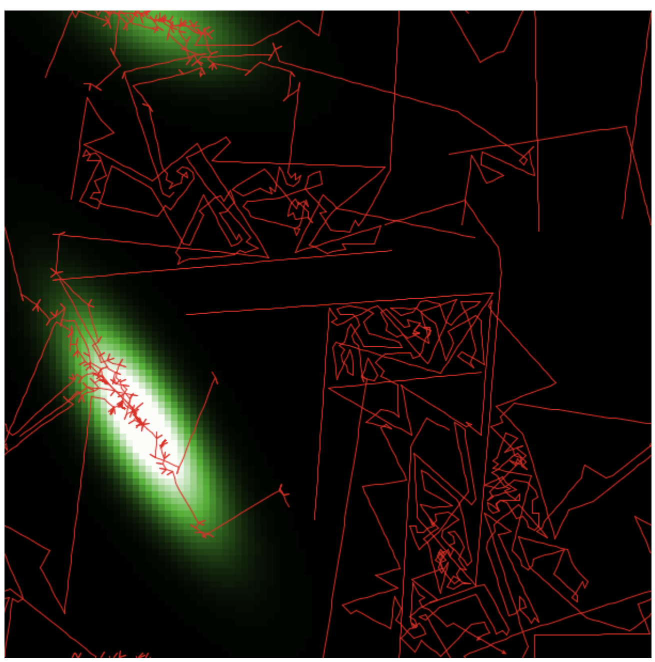


Figure 3.3: An example of maverick behavior in the Muldoon & weisberg model. One can see that the general movement goes straight, but once at a significant patch they will follow other mavericks on the exact same path

But, when mavericks have no unvisited patches in their neighbourhood, they will choose the visited patch with the highest significance. After this, we see mavericks resume their paths, strictly following the agents before them.

This is unreasonable and not very maverick-y. What we implemented is that mavericks do the same thing in the beginning: looking for patches ahead and seeing whether they are less or more significant than their current patch. If it is higher than the previous patch, they keep moving forward; if it is not higher, they will change direction. The mavericks also have a movement and perception range between one and ten fixed before the simulation starts. If something on their path ahead is visited, they will move away from this, thus also changing their direction. Once everything around them is visited, mavericks will find an unvisited patch at a minimal distance to extend their exploration. The following is a detailed description of their decision procedure:

- Is the average significance of the current patch higher than the previous patch?
  - **If yes:** Is there a patch in the current direction that is unvisited yet?
    - **If yes:** Choose one of the patches that is in the current direction
    - **If no:** Choose another unvisited patch that is in the neighbourhood, but in another direction and keep heading towards this direction.
  - **If no:** Choose another patch that is unvisited and closest to the current patch.
- If there is nothing go back and try another direction.

Instead of using the von Neumann neighbourhood, mavericks use the Moore Neighbourhood. Meaning mavericks can reach further if it comes to diagonal movement. This is because agents get dropped with a random heading direction. For example, when they look five patches ahead, they look at the patches with Euclidian distance, meaning this will always be till the border of the Moore neighbourhood. We need to have this in line with each other for technical reasons. It is reasonable for the mavericks to operate in the Moore neighbourhood, considering they are more exploratory and, therefore, can look further than the followers.

Replicators

The mavericks and followers are discussed above. Both agents aim to find unvisited patches and then visit those to expand the epistemic landscape. Both are likely only to see the first drawn score of a patch. Consequently, mavericks and followers are never entirely sure whether the significance they observed is close to the true significance. Here the replicators come into play. These agents may help the mavericks and followers and guide them to the right place, the highest significant patches. Since followers and mavericks can never be sure how close to their drawn score's true significance is, replicators may need to guide them to the right areas. Guidance is especially needed when agents are near significant patches. When patches are made more certain other agents may observe the more certain patches and steer in the right direction.

Replicators are only allowed to land on visited patches. They will pull the lever when they land on a visited patch for an extra time to get another score out of the normal distribution. This way, the patches' lever gets drawn multiple times, collecting more significance values. The significance scores are summed up and divided by the number of visits. This will give us the average significance of a patch. Below, one can find the behaviour rules of the replicator agent:

- collect all the patches that are within the von Neumann neighbourhood of the agent.
- filter these on the fact of being visited and being visited for less or equal to 5 times.
- **Ask:** Is this set empty?
  - **If yes:** Do nothing
  - **If no:** choose one of the available patches with a chance proportional to the average sampled significance.

Replicator agents can, just like the followers and mavericks, execute approaches within their range. We use the von Neumann neighbourhood because this is the most appropriate within our framework. We experiment with ranges going from one to ten. The ranges represent how flexible, knowledgeable and versatile in changing methods an agent is. A range of one is ignorant about their surroundings, and a range of ten is extremely knowledgeable and flexible. For a replicator, it is important to realise that it has to have the right equipment and theory about the approaches already done. Therefore, it can not jump across the grid, but it has to stay close to what it replicated before. The second step is to filter all the patches the replicator can visit. Because it is a replicator, it can only visit already explored patches.

Furthermore, we want to keep the notion of time in mind. Consequently, we want to avoid replicating an approach over and over again. Replicators have a limit of five times replicating an approach. After that, the average significance should be close enough to the true significance to judge the approach well. The only time a patch gets visited more than five times is when the replicator has no options left. It will stay on the same patch, and since agents are called to draw a number every time they are on a patch, they will draw a new number from the current patch they are on. If the agent still has options, for example, one patch with an average significance of 60 (patch A), another with 30 (patch B) and the last one with ten (patch C). These patches will be visited by chance proportionally to their value. For instance, it will visit patch A with a chance of 60 percent. Consequently, replicators will get closer and closer to the patches with high significance and eventually land on the patch with the highest assessed significance. As a result of this, they will reach their goal area.

3.3 Performance Measure & Simulation Runs

We have discussed how the landscape and the distinct agents are modelled. In short, our model has a landscape with patches with a significance score scattered around the true significance of the patch. Once an agent lands on a patch, they will execute the approach, and from this approach, they will get a significance score drawn from the normal distribution. The distribution will have a mean of the true significance of the approach with a standard deviation of 30% of the true significance. Consequently, true higher significance will give a larger variance than true low significance. Significance will be distributed over the landscape by two hills, meaning that higher significance lies close to each other. Followers and mavericks will explore the landscape. Followers, by keeping close to what seems significant approaches and mavericks by moving away from previously visited approaches. Followers and mavericks will visit patches only once, with the exception of sometimes returning to a previously visited patch to alter their direction. The replicators are the agents that will reduce the noise of the drawn significance. Replicators will revisit patches to draw more significance scores for the same patch. Agents will base their decisions on the average sampled significance of each patch. For example, a follower draws a significance of 80. Based on this, a replicator visits the same patch and draws 120. The average that other agents will act on is now a significance of 100. It may be possible that 80 was too low for other agents to explore the area. As a result, agents will move away from this patch. A score of 100 could be enough to draw the agents back to the area.

The agents aim to find the significant peaks as fast as possible. Different distributions of agents will get tested on their performance. On the first runs, a total of 40 agents are used. Then their shares are adjusted each run. This way, homogeneous groups of agent types and heterogeneous groups get examined. The shares are getting increased or decreased by 10%. Each combination gets tested. For example, does it already help to have only 10% of replicators, or do significant differences only show from 30% replicators? The same holds for followers and mavericks. We saw in Muldoon & Weisberg that homogeneous groups of Mavericks were superior to any other group. However, Thoma showed with her adjusted model that a mixed group of followers and mavericks were superior when the movement was non-local.

In addition, influences of different ranges of perception and movement for the agents get tested. All agents will have ranges from one to ten. Mavericks and followers will always have the same range of movement; replicators have their own range of perception and movement. Therefore all the combinations of follower/maverick ranges and replicator ranges will be tested. Thus, to test whether Thoma's outcomes still hold in an uncertain environment. Namely, when movement is non-local, mixed groups are superior compared to homogeneous groups of agents [3]. Moreover, to test from which range replicators have the biggest impact. Lastly, all combinations run for a total of 70 or 100 rounds. Unless there are no mavericks and no followers, then the simulation does not run. Replicators cannot operate independently, so it would not give valuable information to let those simulations run.

We decided to make the landscape more uncertain in the second simulation run. The agents are tested on their behaviour with a standard deviation of 50% of the true significance. This determines whether replicators can let everybody move to the right patches and that other agents will not get lost in sub-optimal areas or agents may not be impacted by the uncertainty and still find their way. All the variables can be seen in table 3.1.

variables	values
total agents	40, 100
proportion agents	0.1 - 1.0
standard deviation	0, 0.3, 0.5
ranges followers	1-10
rounds	70, 100

Table 3.1: An oversight of the variables that are used. Proportions need to add up to 1.0. All possible combinations of followers, mavericks and replicators are being tested. Standard deviation is calculated as the proportion of the true significance. With zero there is no uncertainty, with 0.5 there is a lot of uncertainty. Ranges apply to all the agents. Mavericks and followers always have the same range. Replicators have their separate ranges. For the results shown we used a total of 40 agents, 0.3 standard deviation and 70 rounds. Other parameters were for control runs

The performance of different aspects is tested. Firstly, epistemic growth is measured. Epistemic growth is measured as the proportion of visited patches with true significance higher than zero to that of all the patches with significance. In other words, what percentage of the significant available patches are found? Comparatively, the proportion of the total available significance is measured. It is more important to find the most significant patches, not the whole landscape per se. The lowest significance on the landscape is zero and the highest is 727. When the agents end up with a high proportion, it is most likely they found the significant peaks.

Additionally, because the replicators are trying to get the average significance of the patch closer to the true significance, the difference between the average sampled significance and the true significance of patches is measured. This measurement is done by finding the absolute difference between the perceived significance and the true significance and summing up those differences. Here, one can find out how well the replicators performed. Of course, this number has to be as close to zero as possible. Also, we measure the difference between the top ten found patches. So, collecting the top ten patches, find the absolute difference between the assessed significance and true significance and summing them up to get to the top ten estimation difference.

Furthermore, the patch with the highest assessed average gets compared with the patch with the highest true significance. Ideally, these are the same patches. The distance between these two patches is calculated if they are not the same. This distance is measured in Manhattan distance. Every step going north, east, south or west gets counted until the highest significant patch is reached. Lastly, the total, unfound, and found significance are being tracked. With all these outputs, we hope to find a conclusion on how well all the agents behave. In the following, we will go over the results of the simulations.

Chapter 4

Results

This former discussed all the different components of this model, including how the performance of the model will be measured. We create a baseline to test later how our new agent, the replicator, performs. How do the followers and mavericks perform without the replicator in an uncertain landscape? Four plots are shown. Firstly, what proportions of the significant patches are found? This proportion is called epistemic growth. With this, we will know how much of the valuable patches are explored by our agents. Secondly, what proportion of total significance is found on the landscape? This measure focuses more on the more significant patches. High significance will have more weight in the equation, and a higher proportion tells us that the agents are more likely to be able to find the most significant patches. Thirdly, the absolute difference between the sampled average of the significance and the true significance is called the estimation difference. Followers and mavericks will only sample one value for the patches they visit. From this, we know how much noise is added by the uncertainty. Lastly, the distance between the patch with the highest sampled average significance and the highest true significance is measured. The agents must get as close as possible to the highest true significance. Therefore, the distance between those patches will be measured (in Manhattan distance).

4.1 Model Baseline

fig. 4.1 and fig. 4.2 show that when the number of mavericks increases, the proportion of significant patches and the proportion of total significance that is found decreases. In fig. 4.1, which is about exploring the whole landscape (epistemic growth), one can see four lines. Three of them are selected sample runs based on the agents' flexibility, and one is the average of the complete set of simulations. One can observe that adding four mavericks already lowers the epistemic growth; all the lines go down once the number of mavericks increases. Only with a flexibility range of one, there is no clear finding, and it is subject to much noise. Unless there are more than 50% mavericks, the epistemic growth worsens. In the second graph, which considers the total of the significance and therefore focuses more on the most significant patches, the range of one is also inferior to the others; again, there appears to be no clear finding. The difference with epistemic growth is that the slope is less steep with a higher range of movement, which means that the mavericks are still helping with finding the most significant patches. Although, when there are more than 40% mavericks (16 mavericks), the found significance decreases significantly. In this graph, we also provide three selected samples of distinct flexibility ranges and an average over the full set of simulations

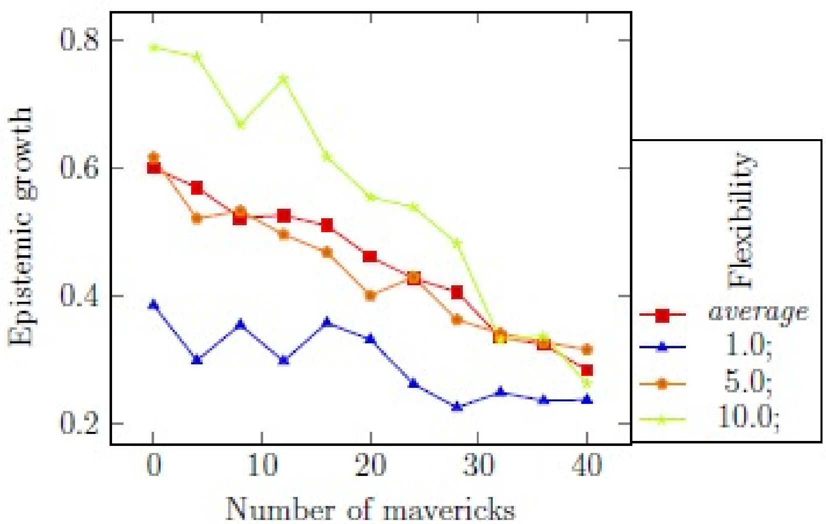


Figure 4.1: Influence of the number of mavericks on the epistemic growth. Three selected samples of flexibility ranges and the average of the full set of simulations. Without any replicators

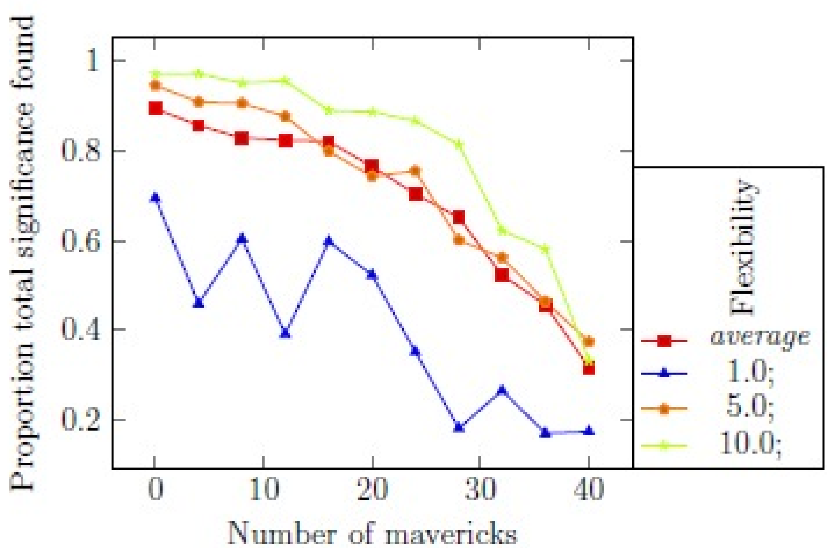


Figure 4.2: Influence of the number of mavericks on the proportion of the total significance found. three selected samples and the average of the total set of simulations. Without any replicators

When one looks at the difference between assessing the patch and the true value in fig. 4.4, no actual pattern is to be recognised. Meaning followers and mavericks recognise the good patches poorly. At the end of the simulation, the group has a significantly wrong picture of the significant patches. fig. 4.4 looks at the top-10 found patches. The top-10 found patches are always the top-10 patches on the landscape. There is also no actual pattern in fig. 4.5 to be found. This graph tells the distance between the best true patch and the highest sampled average patch. The distance varies from around five patches to around 80 patches. In some cases, the high sampled average at the other peak explains the high distance. It is possible that agents assessed the peak at the other side of the grid higher. Therefore the distance is more extensive.

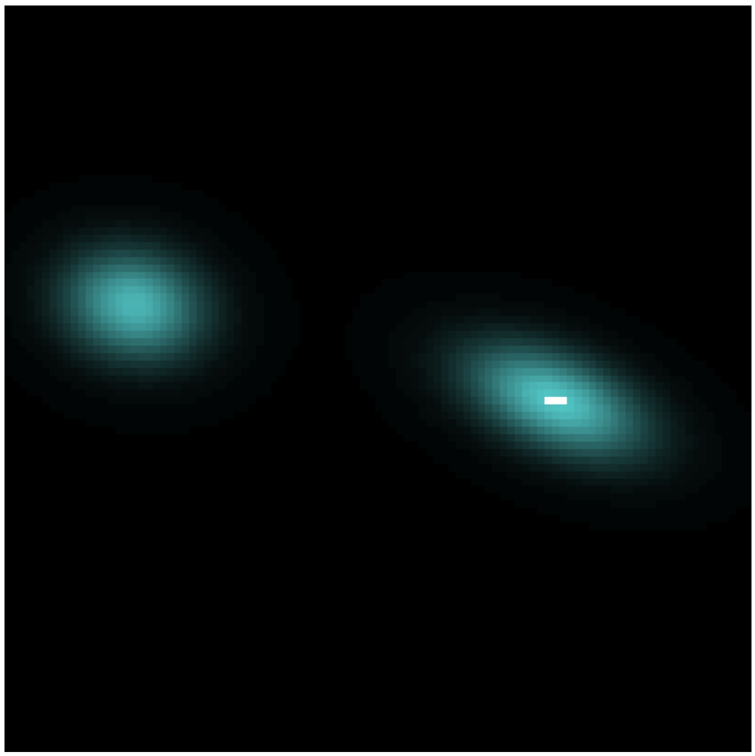


Figure 4.3: An epistemic landscape. The right peak on the landscape consists the patch with the highest significance patch (recognised by the white patches) the left is lower than the right peak.

The left peak is the sub-optimal peak, but the right peak contains the highest value. Although the agents have drawn a higher value at the sub-optimal peak, we can conclude that the agents did not assess the sampled average highest patch as the true highest patch.

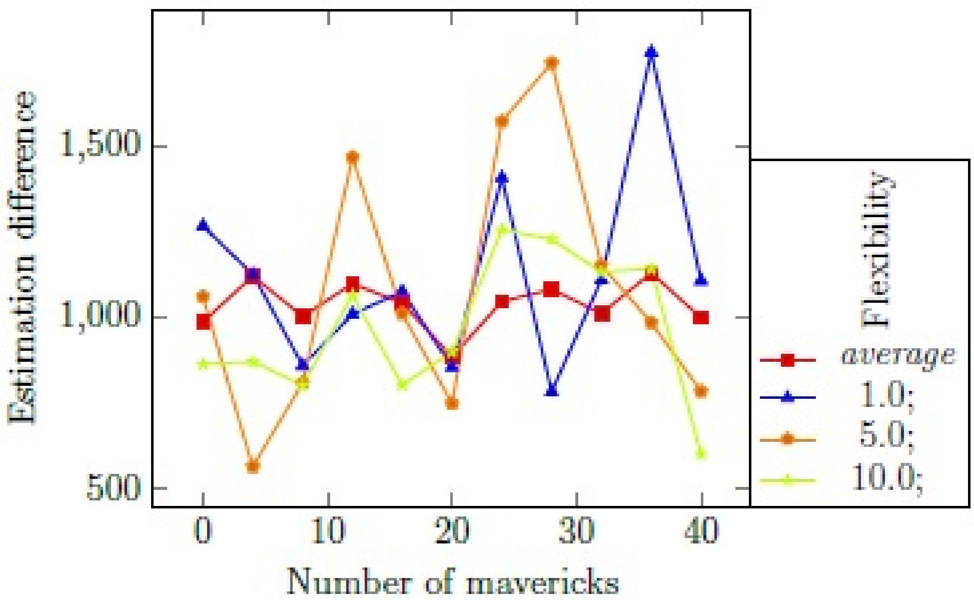


Figure 4.4: Influence of the number of mavericks on the estimation difference for the found top-10 patches. Without any replicators

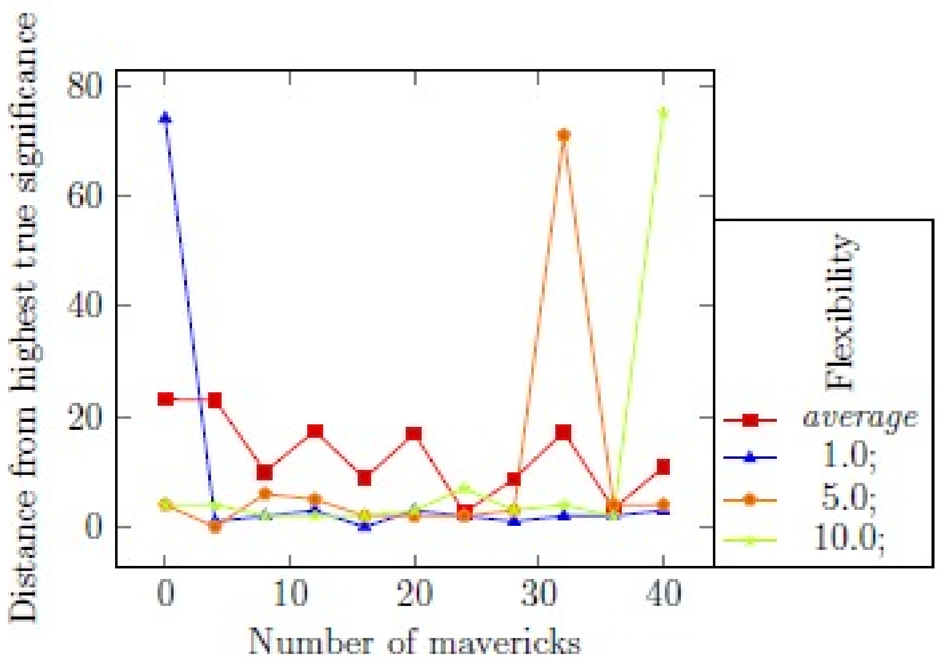


Figure 4.5: Influence of the number of mavericks on the distance between the highest sampled average patch and the highest true significant patch. Without any replicators

4.2 Comparison with Thoma's Results

Our model and Thoma's model produce different outcomes. From her model, she concluded that when movement is local, homogeneous groups of mavericks are superior to heterogeneous groups of mavericks and followers. This result changed when movement was non-local. Heterogeneous groups of mavericks and followers were especially superior when they had a flexibility range of three or higher. Our baseline model does not confirm this. Instead, it shows that having homogeneous groups of followers is always better. With higher ranges of movement as well as with low ranges.

It is beneficial for the agents to have a higher range of movement. This is also the case for Thoma's model. From a range of five, it becomes clear that this is the superior strategy. Of course, one relevant difference between both models was uncertainty. To determine whether this was a decisive factor for mavericks not being helpful, a landscape without uncertainty was tested as well. Here the standard deviation was 0% of the true significance. As a result, agents draw the true significance immediately. Here we find that the outcome is similar to our original findings. The group of followers are still superior to groups including mavericks, and at whatever range, it is beneficial to have a homogeneous group of followers. The landscape that was used for testing was unable to reproduce the same results as Thoma.

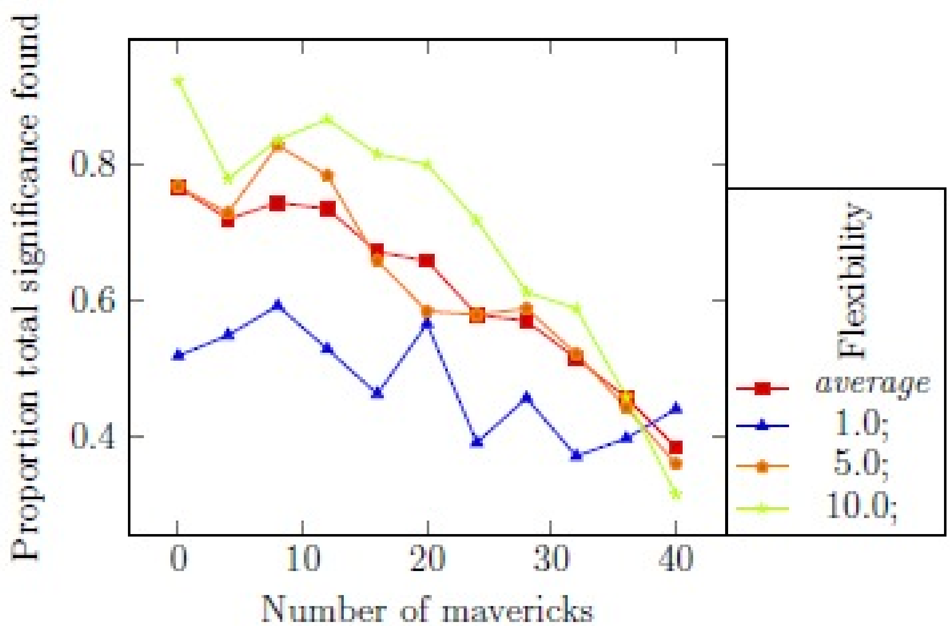


Figure 4.6: Influence of the number of mavericks with the help of replicator on the proportion of total significance found. On this landscape Thoma's results are irreproducible

However, with at least one different landscape, it was possible to replicate her results. The specific landscape has a very high peak at the bottom corner and a lower peak at the right side of the middle. For Thoma, it holds that when there are 50% mavericks, the group performs at their best. Therefore, she argues that mixed groups are superior to homogeneous groups. Almost at all ranges (except for the range of one), it is beneficial to have at least 10% of the 40 agents being a maverick. In the tables of table 4.1 below, one can see the average proportion of significance found. In all these cases, one can see that the average rises between the four and 16 mavericks and then starts to decrease again. With this, we cannot replicate Thoma's results for most of the landscapes. However, it is possible to replicate her results for at least one landscape. Our behaviour rules might be one of the reasons for not replicating Thoma most of the time. Another reason might be that the results are landscape specific. Both reasons will be discussed thoroughly in the following chapter.

Table 4.1: Relation between the number of mavericks and the proportion of significance that is found in a specific certain landscape with distinct ranges

(a) Flexibility range of 5

# mavericks	Average significance found
0	0.54
4	0.61
8	0.67
12	0.50
16	0.57
20	0.48
24	0.48
28	0.42
32	0.39
36	0.25
40	0.29

(b) Flexibility range of 8

# mavericks	Average significance found
0	0.64
4	0.62
8	0.63
12	0.64
16	0.73
20	0.67
24	0.63
28	0.68
32	0.49
36	0.34
40	0.29

4.3 Including the Replicators

In the following, we will analyse the data, including the replicators. To have a positive contribution, it is important that the replicators will not lower the proportion of significance found by a substantial part. Also, we examine whether they lower the estimation differences and the distance from the highest true significant patch and the highest assessed patch.

As shown in fig. 4.7 below, when the number of replicators increases, the proportion of significance that is found decreases. This can be explained by the fact that replicators do not explore. Exploring is the goal of the other agents. In the previous graphs, it is understood that followers are very good at exploring all the significant areas. One thing that does not happen is replicators helping the other agents to the significant patches. Although they are in a uncertain landscape, mavericks and followers are perfectly able to find the peaks on the landscape. Nevertheless, whether or not a significant patch is found does not imply that the agents understand the true significance of the patch. It only implies that the highest true significant patch is visited. The purpose of the replicators will become apparent in the following graph.

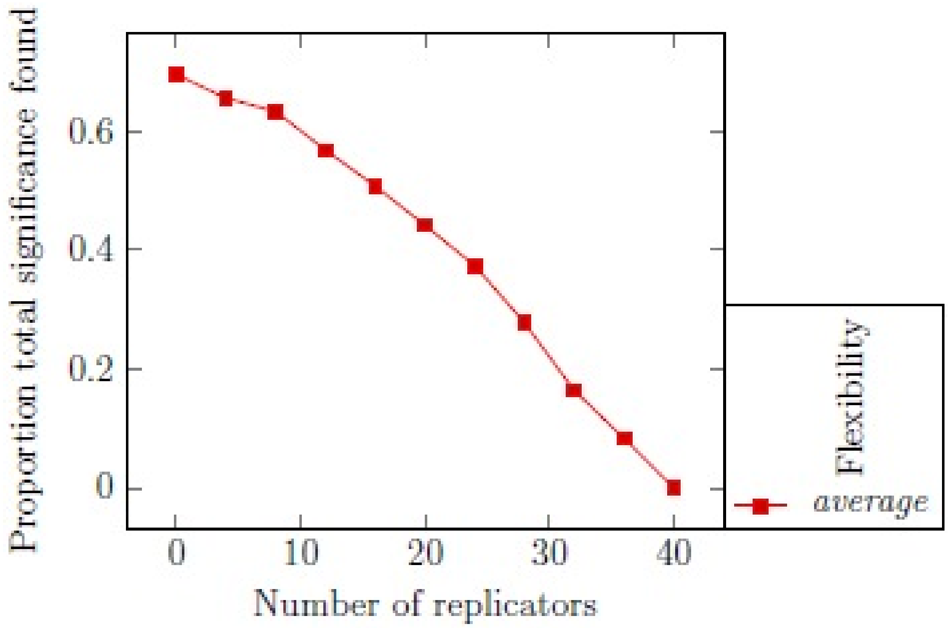


Figure 4.7: Average influence of number of replicators on the proportion of found significance

Once replicators replicate previously visited approaches, one can see that the estimation difference of the top-10 patches gets lower once the number of replicators increases. Where no actual pattern was to be seen without the replicators, one can notice that now, with the increase of replicators, the estimation difference decreases. The average estimation difference for the whole field without replicators was around 33.000, whereas, with replicators, the estimation difference was around 22.000, a significant improvement regarding the judgement of patches. On average, the estimation difference on the top-10 patches was 1017 without replicators and 875 with replicators, which is also an improvement. In this aspect, replicators added epistemic progress. The second thing that stands out is that the estimation difference decreases even faster when the range of movement increases. With a range of ten, the replicators got an average estimation difference of 711. The ability of replicators to perceive highly significant patches at a farther range explains this difference. So, when agents are able to skip patches that are only a slight improvement, they perform better. As a result, the top-10 patches are reached faster than with lower ranges.

Surprisingly, the average distance from the highest patch does not decrease with the help of replicators. The average distance was 16.7 patches without replicators and 19.4 with replicators. However, the distance decreases with low amounts of replicators. The average distance when there were eight replicators was 14.5, which is an improvement compared to those without replicators. The other thing to be observed is that the distance reduces when the range of the replicators is between six and eight, with the best range being a range of 7. The average distance with a range of six was 10.44 patches away from the highest patch.

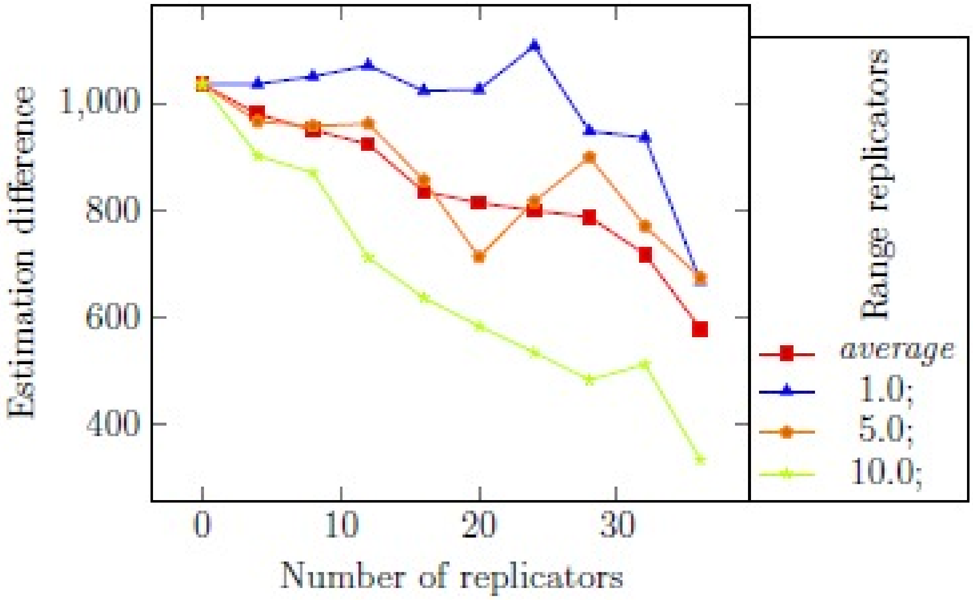


Figure 4.8: Influence of the number of replicators on the estimation difference of the top-10 found patches. Three sampled flexibility ranges and the average of the full set of the simulation

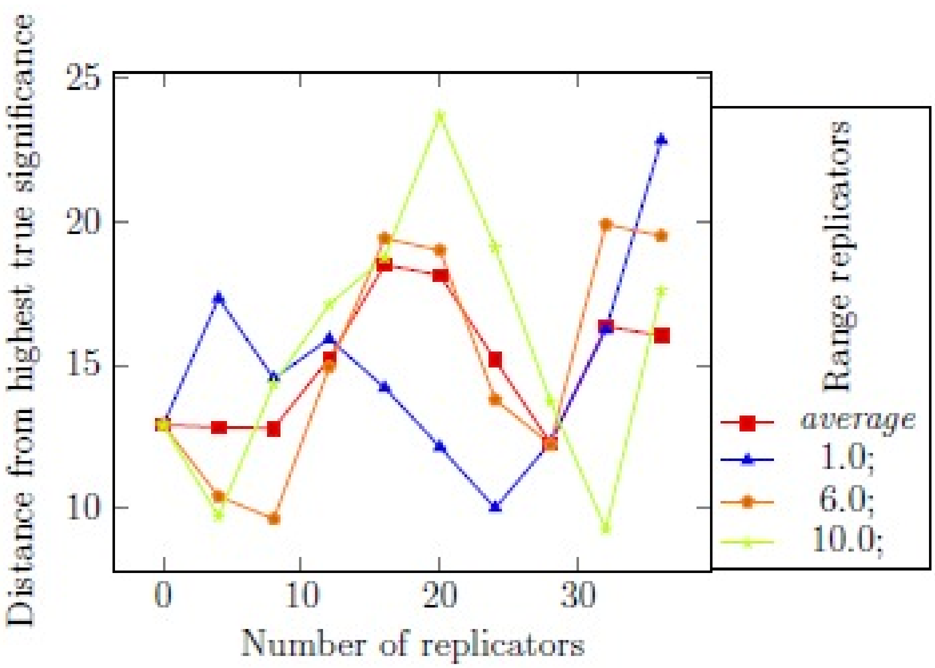


Figure 4.9: Influence of the numbers of replicators on the distance between the highest sampled average patch and the highest true significant patch

Possible explanations may be that replicators try to examine the area more carefully

Table 4.2: Relation between number of replicators and the average distance from the highest assessed average patch and the patch with the true highest significance (left). The range for 8 replicators in relation to the same average distance (right)

(a)

#Replicators	Average distance
0	16.7
4	16.7
8	14.5
12	17.4
16	23.6
20	25.8
24	26.0
28	21.1
32	18.5
36	15.5

(b)

Range for 8 replicators	Average distance
1	15.7
2	16.4
3	17.5
4	16.6
5	14.5
6	11.3
7	10.4
8	15.0
9	13.5
10	13.7

with medium ranges, but they are still flexible enough to skip fewer essential patches. When an extensive range is applied, replicators may replicate patches far from the highest patch. One example could be that, if a replicator has a large range, it has a lot of options to replicate. A patch can have the highest average significance, but there might be 10 other patches that are options too. The replicators has a chance to choose a patch proportionally to their range. It may be the case that chances are higher for choosing the other nine patches, instead of the best patch. There are only so many options for the replicator to replicate when there is a smaller range. Now the chances of being replicated are getting higher. Therefore the assessed average may get corrected more efficiently. However, replicators still need a high enough mobility range to reach the relevant patch soon enough. Looking at the right side of fig. 4.8, the estimation difference is substantially decreased. Though, this may tell us little because when there are more than 90% replicators, much of the landscape will be left unexplored, likely resulting in a low estimation difference.

A final finding is that when flexibility is on the lower side (between two and five), it is better to have more mavericks to have epistemic growth. On the contrary, when flexibility is higher (above five), keeping the number of mavericks on the lower side is beneficial for getting epistemic growth. This difference can be seen in fig. 4.10, where the line above has high flexibility; the two lines go down once the number of mavericks increases. The bottom two lines go in the opposite direction. The flexibility also holds for the followers; this may explain the decrease while flexibility is high. When flexibility is high, followers are very good at exploring the whole landscape. They have big steps and will find significance reasonably quickly. However, when flexibility is low, it takes much longer for followers to explore everything. Here the mavericks come to play; they will go through the landscape much more quicker. Moreover, when they stumble on other visited patches, they will change direction and search for unexplored areas. The followers can follow these explored paths, leading them to significance. In short,

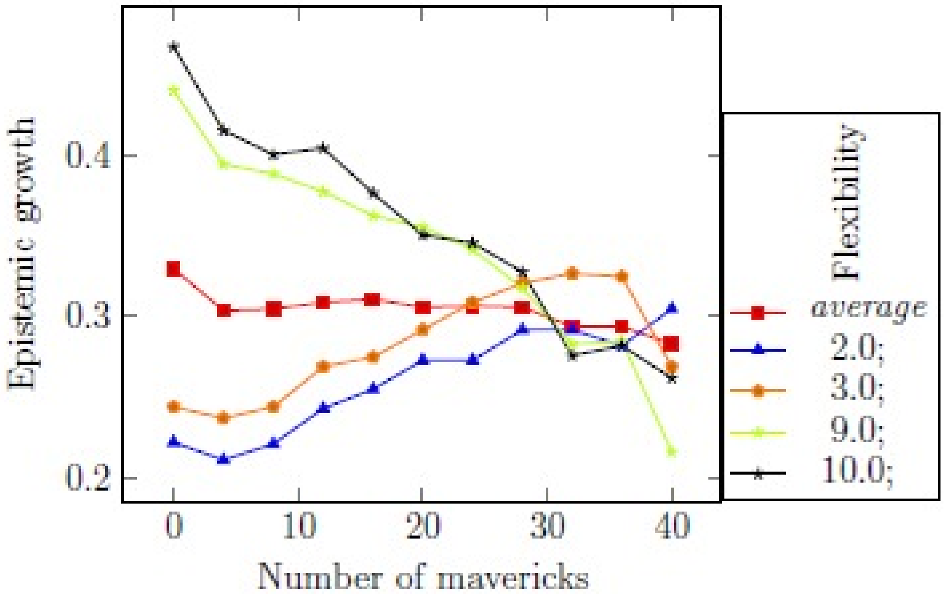


Figure 4.10: Influence of the number of mavericks on the epistemic growth with the number of mavericks. With low ranges more mavericks are helpful. With high ranges homogeneous groups of followers are superior

with less flexibility, the mavericks help the followers by paving the way. With much flexibility, the followers can do this on their own.

4.4 Robustness

For the second set of runs, we have changed the normal distribution's standard deviation to draw a significance value for a patch to 50% of the true significance. Small changes are made to test for robustness. When changes are added to the model and results stay the same, we can judge whether our results are robust. The average assessed significance gets more uncertain, possibly changing the agents' behaviour and increasing the need for replication. However, most of the outcomes stayed qualitatively similar. No real difference exists between the proportion of all the found significance with increased standard deviation. The same holds for the proportion of patches with significance that is found. The estimation difference between the whole landscape and the estimation difference between the top-10 patches also improves with the help of replicators. The difference always lowers with more replicators. Furthermore, the difference lowers when the range extends. Therefore, a range of ten always seems the best to lower the estimation differences. Regarding assessing the best patch, having eight replicators at a range of seven is best. This is also very similar to the results with the smaller standard deviation for the normal distribution.

As a final third set of runs, we examined whether the number of rounds the agents explored was too high. The agents could only explore for 70 instead of 100 rounds during the runs. This did not have any major influence on the results. It did not influence the estimation differences or distance from the best patch. However, the epistemic growth and proportion of found significance decreased slightly. Nevertheless, the agents were able to arrive at the peaks and do their exploring at those areas. This tells us that our results are robust and can tolerate some perturbations in changes in the number of agents, rounds and uncertainty. In the following chapter, we will discuss all our results. Furthermore, the idealisations of the model will be discussed. At last, this thesis will be concluded.

Chapter 5

Discussion

The model results showed the capacities of our distinct agent types; the follower, maverick and replicator. All of them have their function in identifying significance in our uncertain epistemic landscape. The interest lies in their joint endeavour. The goal was to find the proportion of agents needed to find and correctly identify most of the significance on our epistemic landscape. Some of the agents proved to be more helpful in these tasks than others. Our results show that followers and mavericks are efficient significance finders. Finding the hills of significance is no problem for them, even when significance is uncertain. We hypothesised that due to uncertainty, the followers and mavericks would potentially pursue approaches that were of low significance. The uncertainty is caused by drawing a number from a normal distribution. The normal distribution had the true significance as the mean and a proportion of the true significance as the standard deviation. For example, expected situations were where followers and mavericks move the wrong way because they draw a low significance from a high true significance patch. Because agents move in the wrong direction, finding the most significant patches becomes more difficult. However, once some significance was found, followers and mavericks could always upgrade to the most significant patches. This result was even replicated when there were no replicators to help. Arguably, by our results, the replicators are not as functional as expected. They mainly function as minimisers for the estimation differences. They create a better image of the value of all the patches for the agents. Moreover, replicators provide a better understanding of the best approach to the epistemic landscape in some proportions because the replicators reduce the average distance between the highest assessed average significance and the highest true significance.

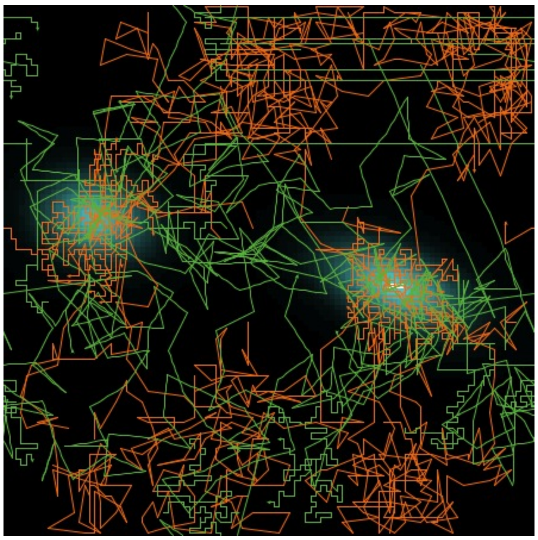
There are still a few points to be made about our model and results. In this section, these points will be discussed. First, we will highlight the idealisations of our model. Secondly, possible reasons why we cannot replicate Thoma's results will be discussed, and finally, some general remarks about formal modelling in the philosophy of science. We will start with why we did not replicate Thoma's results. This will be followed by the idealisations of the model. This chapter will be ended with a discussion about evaluating formal models in general and how this applies to our model. After this chapter, we will conclude this thesis.

5.1 Irreproducible Results Thoma

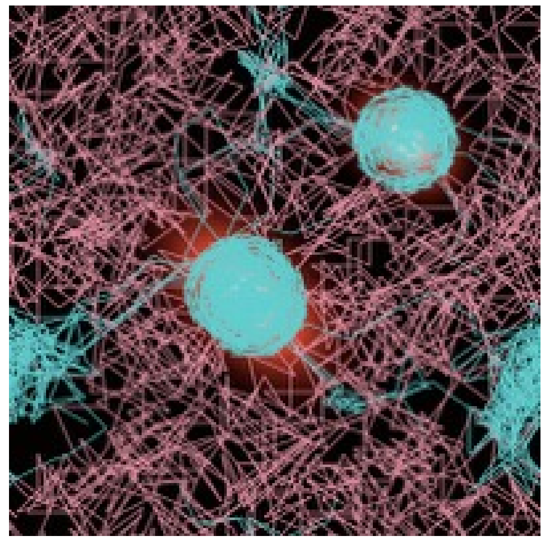
Thoma stated three different results with her model of epistemic landscapes. First, homogeneous groups of mavericks are superior to others when local movement occurs. In other words, when mavericks and followers were only allowed to move one patch at a time, a group of only mavericks was superior to others. This result is similar to the work of Muldoon & Weisberg [1]. However, when agents were more flexible and knowledgeable about their surroundings, having mixed groups of mavericks and followers became beneficial. Whenever agents were allowed to take steps of more than three patches, it became clear that mixed groups were superior. Finally, when movement was global (travel wherever they want), the result was still the same, but the groups tended to take longer to converge to the highest patch. However, our model could not produce the same results as hers on certain or uncertain landscapes. The results that were not replicated were the superiority of groups of mere mavericks with local movement and the superiority of mixed groups of followers and mavericks with non-local movement. We were only able to replicate her results on one specific landscape, which we will discuss later in this section.

Possible reasons for this are: firstly, our follower agents have slightly different behaviour rules. Compared to our followers, one can see that her followers always travel the same amount of patches according to their flexibility range. When her agents have a range of six, they seem to always skip six patches. In contrast, our agents travel within a neighbourhood. With a range of six, they will look at patches within their von Neumann neighbourhood with a range of six. Consequently, our followers are allowed to take steps between one and six patches if that is beneficial. The difference is shown in fig. 5.1. One can see that the agents behave differently, especially in the vicinity of the peaks. In Thoma, it may be that Thoma's movement rules force agents to travel too far. Even if it would be beneficial to take a step of, for example, four, the rule forces the agents to take a step of six in this case. Our agents, in contrast, would have the option to take a step of four if this is beneficial for finding the best patch. This can be seen in fig. 5.1. Our followers (in green) take smaller steps when they are close to or already on significant patches. They likely take small steps because there are visited patches with high significance in their neighbourhood, and the rules force them to visit neighbours of highly significant patches.

Secondly, it stands out that the significant areas are very densely visited in Thoma's model. Her agents may be walking in circles around the hills, exploring the hills extensively and thereby creating these results one can see in fig. 5.1. However, the agents likely revisit patches that have been discovered before. Otherwise, the paths would stand farther from one other as they do now. In our model, this is prohibited behaviour. Consequently, our followers also expand their knowledge outside the peaks because other patches are still optional because they are not visited yet. This might be why, on average, our agents find more significance per simulation run. One could argue that in our model, there are visible holes in the peak, and Thoma's peaks do not have any holes. This possibly means that Thoma's agents found many more significant patches than our agents. However, the data does not support this argument. In fig. 4.2 on page 33, one can see that our agents in most cases, reach a proportion above 0,8 of total significance found in 100 rounds, whereas Thoma's agents never surpass a proportion of 0,8 total significance found [3]. Our agents will keep exploring while it may be possible that the best patch is already found. Our agents do not know whether the best patch is found if patches are still unvisited because unexplored patches can still have high significance. Therefore, they will keep exploring the landscape. The movement restriction on Thoma's agents and revisiting explored patches may cause the inability to replicate Thoma's results.



(a) Our model. In green the mavericks and orange the followers



(b) Thoma's model, figure 7 of epistemic division of labour revisited[3]. In pink the maverick and light blue the followers

Figure 5.1: The difference between our model simulation and that of Thoma.

However, on one specific landscape, it was possible to duplicate Thoma's results regarding the superiority of mixed groups in finding the highest proportion of significance. In fig. 5.2, one can see the specific landscape we found the results. There is a high peak of 1832 significance in the left-down corner and a lower peak of 500 significance right next to the middle. The simulations showed that without the help of mavericks, the left down peak did not get visited a lot. Resulting in lower proportions of total significance found. The left peak was visited more frequently when there were more mavericks. The best results were found when there were between 4 and 16 mavericks. With these proportions of agents, the left peak was visited by mavericks and followers. This may indicate that followers were drawn with the mavericks to find the significant patches. When there were no mavericks, the sub-optimal peak was visited more often than the optimal peak. Between the four and sixteen mavericks, the groups performed at best. From this, it can be concluded that the results of our model are landscape specific. Either the different heights of the peaks, the locations of the peaks, or a combination of the two cause this outcome. This should be investigated in a more detailed manner to verify this result. In the following sections, we will discuss our model's idealisations and limitations. In the last section of this chapter, we will discuss the process of evaluating formal models in general.

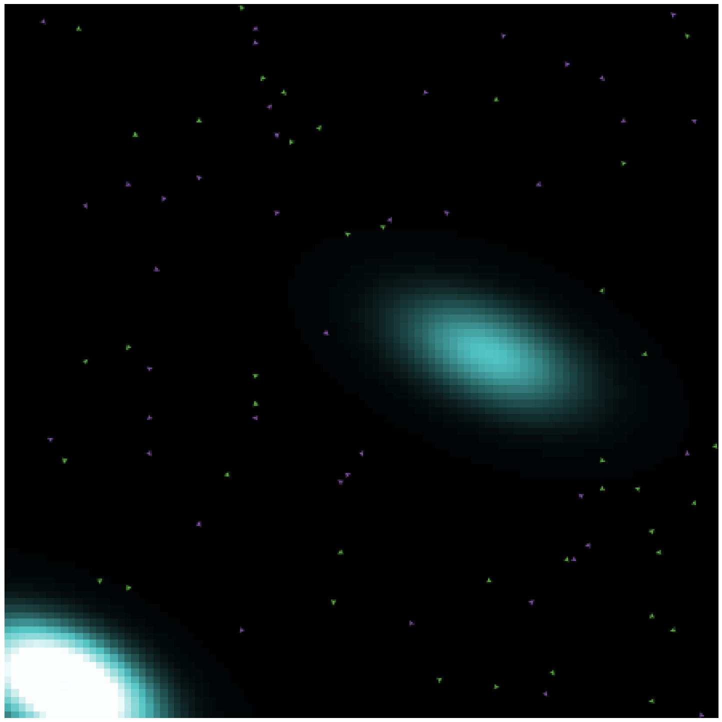


Figure 5.2: The epistemic landscape where we were able to replicate Thoma's results

5.2 Model Idealisations

Our model has various idealisations. These idealisations play a role in the outcomes of the model. These will be discussed in three parts: Landscape, agents, and simulation runs idealisations. Moreover, we discuss the robustness of our model. In other words: with different assumptions, does our model lead to the same outcomes? If this is the case, we can make inferences from the model to the real world [24].

5.2.1 Landscape

Firstly, the landscape specifics of our model will be discussed. We have modelled our landscape as Muldoon & Weisberg did—an epistemic landscape with two peaks of significance generated by a Gaussian formula. The formula places the peaks at random positions and with random heights. The heights of the two peaks can be similar or differ a lot from each other. In our case, the peaks were almost the same heights, although their shapes differed. One of the peaks had a slope that was less steep than the other. Our highest peak had a maximum significance of 500, and the highest significance of the other peak was 457. Furthermore, our peaks are relatively close to each other. There are a few options to explore left. One could also test other distributions of peaks. It is not always the case that there are two good answers to a research problem. Sometimes it would be the case that there is only one answer, and it may also be the case that multiple answers are available in some situations—for example, the treatment of heart disease. A variety of medications is used to treat heart disease. Some medications include blood thinners, statins (to reduce cholesterol) and Beta-blockers (to lower heart rates). Sometimes only one of the medicines is enough, but there may be cases that a combination is needed. Further research should investigate whether different topologies of peaks influence our model's outcomes.

Moreover, one could test other distributions of height. We have seen that on most landscapes, it was not possible to duplicate Thoma's results. Namely, it was not beneficial for our model to have mavericks in the groups of scientists to find the most significance. In contrast, Thoma's mixed groups of agents performed best when movement was non-local. However, on one landscape, mixed groups of agents were superior to homogeneous groups of followers. This is discussed in the previous section. One of the differences was the height of the peaks. One of the peaks had a height of 1832, and the other peak only had a height of 500. This might be an important factor for the ideal distribution of agents. In our case, we saw that the highest significance was only visited if there were some mavericks. When the mavericks found the highest peak, the followers could follow the mavericks to the highest significance. The other factor could be the locations of the peaks. Although, the landscape does not have edges, so once an agent goes too far to the left, it will end up on the right side of the landscape. The same happens with up and down. This has as consequence that the landscape could always be rotated in some way that it will look very similar to other landscapes regarding the locations of peaks.

The following point is about the shape of the hills. On our landscape, the highest significance of the hill is in the middle, and significance gradually decreases around the peak. There are a few options to explore left. One of them would be to test narrower and wider peaks. Wider peaks represent the possibility of more combinations of approaches to work. One may recall the description of an approach by Muldoon & Weisberg: the combination of background theory, measurement devices, output analysis and research question. It could be that an exact combination of the four is needed to find the perfect solution. For this, a very narrow hill is needed on the landscape. However, it could also be the case that more combinations offer the highest significance. In this situation, one might model a wider hill. It may also be necessary to look at the shapes of the hills. Another idea would be to model the significance as plateaus. In other words, Only a few very approaches have high significance, and around them, approaches have zero significance. We need to explore these options of adjusting the assumptions of the landscape slightly to judge the robustness of our results. Therefore, more research is needed regarding the shape and height of the hills on the epistemic landscape.

Different degrees of uncertainty were tested. When agents land on a patch, one may recall that they draw significance from the normal distribution. The normal distribution uses the true significance as the mean and the proportion of the true significance as the standard deviation. Standard deviations of 0%, 30% and 50% were used. The former being not uncertain at all and the latter being the most uncertain. All the uncertainty degrees result in the same outcome. Namely, having only followers is better than mixing groups of followers and mavericks. For estimation differences, we also found the same qualitative results for all uncertainty degrees; more replicators help to lower the estimation differences.

For our robustness analysis, we find that uncertainty perturbations do not influence our simulations' outcomes. On most landscapes, the model shows the same results regarding estimation differences, epistemic growth and proportions of total found significance. However, with high uncertainty degrees, the improvement in estimation differences is better with more replicators. On at least one landscape, our results were different. It was beneficial for epistemic growth to have mixed groups of mavericks and followers on this landscape. This may imply that our results are landscape specific and that one could investigate the possibility that there are adequate types of landscapes for specific research problems.

5.2.2 Agents

In our model, each agent operated simultaneously. Consequently, the replicators are always behind the followers and mavericks. Intuitively, this sounds plausible; replicators are always behind because they cannot replicate what has not been examined before. However, waiting for a follower or maverick to see how significant an approach truly is may sometimes be worth it. This may influence the decisions of the follower and maverick. One option is to let agents take turns in exploring the landscape. For example, to let mavericks and followers operate in even rounds and replicators in uneven rounds. With this, agents may be less influenced by uncertainty, and replicators have more influence on the outcome. Some patches may seem insignificant, but after replicating, they can become more significant. Possibly, that patch becomes more significant than another patch in the neighbourhood. Consequently, a follower or maverick may get drawn to another patch than the round before.

With high flexibility, the follower agent type is powerful in finding significant patches. With lower flexibility, mavericks are still helpful in finding most of the total significance. With epistemic growth, having more mavericks when flexibility ranges are low is beneficial. Arguably, mavericks help open up paths for followers with low flexibility, and with high flexibility, this is no longer necessary. Then followers can explore enough of the landscape. Overall, followers, on their own, are excellent significance finders. It is possible that we correctly modelled the follower and that a group of followers is superior to other groups. Everybody incrementally expands their regions; eventually, some will find the most significant approaches. This process gets even more accelerated with increasing flexibility.

Alternatively, the follower rules were modelled so that they were too strong. For instance, once a follower has no available patches left in their neighbourhood, the follower can travel to a patch out of their range. Their movement rules state: if there is no unvisited patch in the von Neumann neighbourhood, go to the nearest unvisited patch. As a result, followers can always find a patch unless the entire landscape is visited. An option can be to let agents get stuck when their neighbourhood is fully visited. Another option is that agents have to skip a round to extend their flexibility range. Waiting can be seen as making oneself more knowledgeable or collecting the required devices. This could make the follower agent less powerful. It is also reasonable to think that a scientist can get stuck momentarily. One could simply run out of ideas or out of money. As a result, mavericks and replicators may be of more importance. Decisions carry more weight when there is the possibility of getting stuck. Mavericks can show the paths to pursue, and followers can help show the true significance of approaches.

The next idealisation is that mavericks and followers always have the same flexibility range. One can argue that mavericks and followers can have different flexibility. It would be interesting to find out whether mavericks are more useful when they are more flexible than the followers. Intuitively, having more flexible mavericks than the followers makes sense because they are already the more exploring type of agent. Future research may find different results when agent types have their own flexibility range. It is possible that only with very small ranges for the followers and large ranges for the mavericks, mavericks may have some more influence. Once followers have a flexibility of more than four, they start to become strong significance finders. However, our results are robust regarding perturbations in flexibility. We can see that high flexibility is always superior to low flexibility. Especially when followers increase their flexibility, they become powerful.

5.2.3 Simulation Runs

The following concerns the simulation runs. The simulations were run on one specific landscape. In NetLogo, there is the option to add a randomiser, which keeps the landscape the same every time a new simulation runs. The landscapes are kept the landscape the same to compare possible influences of other factors. This way, the influences of the flexibility ranges and agent types can be tracked more efficiently. However, one can notice that only some of our graphs have smooth lines across the x- and y-axis. The lines in our graphs are not always curved as some of the lines in fig. 4.2 on page 33, but some lines are subject to a lot of noise. This can also be seen in fig. 4.9 on page 39. In this figure, the blue line indicates that adding a few replicators increases the distance between the true best patch and the best-assessed patch. However, when more replicators are added, the distance decreases. Adding even more, it is then again worse regarding the distance. There is no actual pattern to be found. This may indicate that we need to collect more data points to create an adequate image of the situation. The same holds for some of the other graphs. Collecting more data points might smooth these lines out to create a better image of reality.

We have tested with 100 agents and with 40 agents. Furthermore, we let the simulations run for 100 and 70 rounds. This was all done on the same landscape. Our results stayed qualitatively the same. Minor changes in the number of agents and rounds run did not impact the results. Regarding the number of agents and rounds, our results are robust.

Although, by adding uncertainty, the model is more realistic compared to Muldoon & Weisberg, and Thoma, our model is still highly idealised and could be expanded in several ways. Another expansion one can think of is the distinct capabilities of agents. Not every scientist has the same toolbox. Funding can differ, but also talent and flexibility. Arguably, when a scientist or a group of scientists have more money to invest, it is possible to afford bigger steps on the landscape. Also, some scientists have a better idea of what would be the ideal solution for a problem. We can represent this by giving agents different sizes of neighbourhoods. It is possible to assign different ranges of movement to the agents and examine the outcomes of the different distributions.

Giving agents different ranges of flexibility adds to the complexity of the model.

Further research is also needed to determine the effects of the landscape on finding and correctly identifying significant patches in an epistemic landscape. For example, one could examine whether different heights of peaks influence the outcomes. Moreover, one can test whether the shape of the hills has an effect on identifying significance. At last, one could model a landscape in more than two dimensions. Approaches in the landscape are very fine-grained; a landscape with multiple dimensions may prove to be very useful.

We would want to know how robust our results are with the proposed perturbations and changes of assumptions in this chapter to make more inferences from the model to the real world. We have shown that our model is resistant to most slight changes. We examined changes in uncertainty, number of agents, number of rounds and distinct landscapes. On most of the changes, our results stayed the same. Other assumptions and perturbations need to be tested in further research to discover the robustness of our outcomes. In the following section, the evaluation of formal models is discussed.

5.3 Evaluating Formal Models

5.3.1 Representational and Predictive Accuracy

This chapter will be ended with a discussion of what can or cannot be done with agent-based models and how these models need to be evaluated. An argument against our model might be the absence of a link between real-world data and the outcomes of our model. As a result of this, it loses some of its predictive value. Furthermore, making inferences from our model to the real world may be more difficult. Data could be one way to bridge the gap between the model and the target system. However, sometimes it may be impossible to use real-life data. Nevertheless, this is not a reason to abandon a formal model.

One train of thought comes from Micheal Thicke, who argues that formal models of science ought to be evaluated to the same standards as philosophers set for models in other disciplines [25]. In his paper, Thicke argues that models can be evaluated according to both their predictive and representational accuracy. Representational accuracy is the ability to link the model to the target system. Thicke illustrates this with the famous model of Thomas Schelling concerning racial segregation [26]. Schelling constructed a set of imaginary cities. In these cities, racial segregation evolves when people have preferences about the racial mix of their neighbours. It turns out that strong segregation already happens when people have mild preferences. Thicke argues that it is essential that the model has similarities to its target. In Schelling's model, cities were constructed out of multiple pennies. One penny represents one person, and multiple clustered pennies represent a city. The persons behave similarly to actual people. In a city, people have neighbours, and people can have preferences about their neighbours [25]. When a model has more similarities to its target, it has more representative accuracy, which gives us more reasons to accept its claims.

The second form of accuracy is predictive accuracy. This involves comparing predictions generated by the model with observations of its target system. For instance, we might assess the predictive accuracy of Schelling's model when we compare the boundaries between segregated neighbourhoods in his model and check whether boundaries are shifting in actual cities. If there are shifting boundaries, it could increase the credibility of other model predictions. An example of a prediction might be: ethnically-diverse cities will likely become increasingly segregated when people have only mild preferences for the ethnicity of their neighbours [26]. In other words, outcomes from a model should be compared to empirical data. When a model lacks the capability to have predictive and representational accuracy, Thicke argues that these models should be handled with caution. In his view, with some exceptions, the current generation of formal models cannot support any normative conclusion about science. Therefore, they cannot be invoked in a policy context. If modellers cannot develop some method for evaluating their claims, Thicke does not see reason to take these models seriously.

5.3.2 Other Uses of Formal Models

However, representational and predictive accuracy is only sometimes the purpose of a formal model. Arguably, there are many purposes of formal models that do not need validation. For instance, when illustrating that some events or situations are plausible (sometimes in a meta-physical sense), it is often unnecessary to validate the model [27]. But, of course, models need to be validated in some cases. COVID-19 is a great example. There are many models designed to predict future outbreaks of the COVID-19 virus [28]. These models were used in policy contexts. Think of a lockdown, curfew regulations, and not being allowed to stand within one and a half meters of each other. These decisions were made partly due to the mathematical models. Nevertheless, even in this case, validating the models was not always possible. Often, in the beginning, the models were wrong. Nevertheless, the disease spread fast, and decisions needed to be made.

In other situations, validation is complicated or impossible. Nonetheless, it is not always a reason to abandon these models. Mayo-Wilson argues that simulations for philosophy can be used for two reasons. First, simulations should be used for the same reason we currently use thought experiments. Most of the time, simulations are superior to thought experiments. In a complex social system, losing sight of the bigger picture with a thought experiment is easy. Simulations allow us to model these complex interactions to get a grasp of the social system at hand. With a model, one isolates the important variables in explaining and visualising a social phenomenon. Secondly, devising and coding models help with philosophical habits of mind. To make assumptions is to think about a phenomenon. When modelling a complex system, one has to understand the system in great detail. This helps us create a better image of the complex system for others, but mostly ourselves [27].

Thought experiments can be used for various reasons. Mayo-Wilson states a few: exploring the dynamics of social and physical systems, illustrating conceptual possibilities and exploring logical relationships among philosophical theses. For example, Kant used the example of promise-keeping in his famous Groundwork of the Metaphysics of Morals, and the categorical imperative [29]. For the sake of this discussion, we will not dive into great detail of the categorical imperative. In short, Kant asks us to imagine how people would react to norms changing. In this case, promise-keeping. One needs to fast forward to a situation where no one keeps a promise, which makes making a promise useless. When no one holds a promise, no one will believe a promise. This gives us a contradiction, telling us that breaking a promise should be avoided.

Mayo-Wilson argues that when answering a philosophical question requires understanding the dynamics of social systems, simulations are better than thought experiments. This is because social systems are often complex. Simulations can be used to track the interactions of thousands of agents. All these agents can have different features which can be tested. Models can be manipulable; details can be changed to test the robustness of a conclusion. Moreover, the results of a model can be visualizable, which can help understand the target system in more detail.

The second reason a philosopher should model is that it can help with philosophical habits of mind. When coding a model, one has to translate a target system to a fitting model. One needs to make some assumptions to create a working model. Think again about the example from Kant and promise keeping. If one wants to model this example, one is forced to ask themselves more questions than Kant did in the past. For example, if one wants to model belief, should the modeller model this as binary (is someone reliable or not) or in scaled items (is someone very reliable, somewhat reliable or not at all)? The result depends greatly on assumptions like these [27]. These questions help the modeller create a better image of the situation. Summarised, modelling can help the modeller understand the system target more elaborately. Therefore, formal models have more purposes than representative and predictive accuracy, and we should not abandon a model if these qualities are lacking.

Although Mayo-Wilson argues that evaluation is not always needed. Weisberg offers another way to evaluate formal models [24]. One may validate their model differently when representative and predictive accuracy is lacking. Namely, by robustness analysis [24]. Weisberg describes the robustness analysis as a four-step procedure. It starts with collecting a set of models that try to predict a common result. He calls this the robust property. In our case Muldoon & Weisberg's, Thoma's and our model try to predict a common result. The second step involves analysing the models for a common structure that generates the robust property. This can be the landscape setup or the agents' behaviour in the models. The third step gives an empirical interpretation. With the model, we want to say something about the real world. So, how do we translate our outcome to real-world phenomena? The final step is a stability analysis. For instance, Thoma added flexibility to Muldoon & Weisberg's model. Then, we added uncertainty and another agent. By adding complexity and perturbations, one can further test the robustness of the results, as the results should stay the same to be robust.

More research is needed to test the robustness of our results. In our simulation runs, some perturbations were added, such as changing the landscape, altering uncertainty, and varying the number of agents and rounds. We found that, in most cases, our results were qualitatively similar. However, our results changed when we changed the landscape to the landscape of fig. 5.2.

In this section, we have discussed why we were unable to replicate Thoma's results most of the time. Secondly, we have discussed the idealisations of our model and how we could change the assumptions of our model in the future to test for robustness. With this, we have made some suggestions for future research. This included how our agents behave and how our landscape is modelled. Lastly, we have discussed how to evaluate formal models. One can do this by testing representative and predictive accuracy. However, sometimes it is impossible to find these two qualities. In this case, one may validate their model by robustness analysis. Furthermore, formal models force us to think about social and physical dynamics and can help us to become better philosophers of science. Finally, in the following chapter, conclusions will be drawn from our model and inferences to the real world will be discussed.

Chapter 6

Conclusion

In this thesis, we presented our agent-based model on the division of cognitive labour. This model is an extension of the models from Muldoon & Weisberg, and Thoma. In these models of epistemic landscapes, they examined the ideal distribution of two agent types: followers and mavericks. However, these models are highly idealised and do not account for uncertainty in scientific research. Research can be erroneous, and therefore replication is needed to verify outcomes. The model presented in this thesis aims to include this uncertainty and examines how relevant replicators are in the division of labour. The ideal distribution of agent types was examined to find and identify the most significant patches on the epistemic landscape. Three agent types were used: followers, mavericks and replicators. All had their rationale. Followers represent conservative scientists. They incrementally expand their knowledge by staying close to successfully done research and extracting all the information there is to know in their area. Mavericks are adventurous, paving new paths for others to follow or avoid, depending on what they find. Finally, replicators make the outcomes of previously done research less noisy.

We have three main findings. Firstly, to find the most significant approaches, it is best to have homogeneous groups of followers. Followers are able to detect and exploit approaches that are very useful efficiently. When the agents have more knowledge about their surroundings and can adjust their methods more easily, they become even more efficient in finding the best approaches for the problem. Secondly, there is a difference between finding the best approaches and finding all the approaches that can have value. For the latter, when scientists cannot adjust quickly and do not know much about their surroundings, it is good to have mixed groups of agents. When scientists can adjust quickly, only having followers is more beneficial. Thirdly, it is not only important to find the approaches; it is also essential to correctly label the approaches as they truly are. We found that adding replicators helped assess the approaches' quality correctly. Replicators do also benefit from being more adjustable and knowledgeable about their surroundings. However, with too many replicators, not enough approaches can be explored. As a result, agents miss out on much knowledge and unnecessarily replicate approaches.

It is important to find a balance between finding approaches that are valuable and correctly assessing all the approaches on their quality. The model shows that having many followers benefits the former, and having more replicators benefits the latter. Arguably, it would be best for the scientific community that scientists explore as much as possible and that only the most promising results are replicated. Therefore, the best division of labour should be mostly followers with some replicators added. Enough valuable approaches are discovered with this division, and the most important discoveries can be replicated. As a result, agents will be able to better choose and pursue the best approaches.

Due to the uncertainty, it was expected that followers and mavericks might move in the wrong direction. Thus, the agents move to areas on the landscape representing insignificant outcomes. Replicators were expected to help the other agents show which patches were significant and which were not. However, the model shows that agents efficiently found significant patches, particularly the followers. In contrast with what was expected, the uncertainty of the landscape did not play an essential role in this. For identifying significance, the expectations were met. Once more replicators were added, groups of agents were better able to identify the approaches to their true significance. This means that replicators help reduce the noise of the significance of scientific approaches.

The model presented in this thesis shows different results than Muldoon & Weisberg and Thoma. Muldoon & Weisberg showed that homogeneous groups of mavericks are superior to other mixed groups or homogeneous groups of followers. Thoma's model had the same result for well-informed agents. Although, mixed groups were better when agents were well-informed and flexible in adjusting their research. Except for one landscape, this research could not replicate their results. Further research may determine the influence of the landscapes on the results.

Although replicators in our model are not needed to guide the other agents to the helpful approaches for a research problem. They do create a better image of the actual significance of all the discovered approaches. Even in small proportions, replicators proved to be helpful in this task. As normal science progresses, errors inevitably will occur and may result in pursuing inferior approaches. Replication might prevent this from happening. Moreover, as in the example of Freedberg and Barron, a theory might be one replication away from becoming an accepted truth. [1]

[1] If you would like to receive the code of the model, contact lamarkiel@gmail.com

Bibliography

[1] M. Weisberg and R. Muldoon, "Epistemic landscapes and the division of cognitive labor," *Philosophy of science*, vol. 76, no. 2, pp. 225–252, 2009.

[2] K. J. Zollman, "The epistemic benefit of transient diversity," *Erkenntnis*, vol. 72, no. 1, pp. 17–35, 2010.

[3] J. Thoma, "The epistemic division of labor revisited," *Philosophy of Science*, vol. 82, no. 3, pp. 454–472, 2015.

[4] P. Kitcher, "The division of cognitive labor," *The journal of philosophy*, vol. 87, no. 1, pp. 5–22, 1990.

[5] T. S. Kuhn, *The structure of scientific revolutions*, vol. 111. Chicago University of Chicago Press, 1970.

[6] M. Baker, "1,500 scientists lift the lid on reproducibility," *Nature*, vol. 533, no. 7604, 2016.

[7] C. Martini and M. Fernández Pinto, "Modeling the social organization of science," *European Journal for Philosophy of Science*, vol. 7, no. 2, pp. 221–238, 2017.

[8] J. Scott, "Rational choice theory," *Understanding contemporary society: Theories of the present*, vol. 129, pp. 126–138, 2000.

[9] D. Klein, J. Marx, and K. Fischbach, "Agent-based modeling in social science, history, and philosophy. an introduction," *Historical Social Research/Historische Sozialforschung*, vol. 43, no. 1 (163, pp. 7–27, 2018.

[10] S. Tisue and U. Wilensky, "Netlogo: A simple environment for modeling complexity," in *International conference on complex systems*, vol. 21, pp. 16–21, Boston, MA, 2004.

[11] F. Mali, L. Kronegger, P. Doreian, and A. Ferligoj, "Dynamic scientific co-authorship networks," in *Models of science dynamics*, pp. 195–232, Springer, 2012.

[12] I. Lamprecht, "Calorimetry and thermodynamics of living systems," *Thermochimica Acta*, vol. 405, no. 1, pp. 1–13, 2003.

[13] J. Lovett, M. Ryuntyu, and D. L. Liu, "Allelopathy, chemical communication, and plant defense," *Journal of Chemical Ecology*, vol. 15, pp. 1193–1202, 1989.

[14] M. Hu, Y. Zhong, S. Xie, H. Lv, and Z. Lv, "Fuzzy system based medical image processing for brain disease prediction," *Frontiers in Neuroscience*, vol. 15, p. 714318, 2021.

[15] B. M. Radomski, D. Šešelja, and K. Naumann, "Rethinking the history of peptic ulcer disease and its relevance for network epistemology," *History and Philosophy of the Life Sciences*, vol. 43, no. 4, pp. 1–23, 2021.

[16] D. Šešelja and C. Straßer, "Heuristic reevaluation of the bacterial hypothesis of peptic ulcer disease in the 1950s," *Acta biotheoretica*, vol. 62, no. 4, pp. 429–454, 2014.

[17] P. Malfertheiner, F. K. Chan, and K. E. McColl, "Peptic ulcer disease," *The lancet*, vol. 374, no. 9699, pp. 1449–1461, 2009.

[18] E. D. Palmer, "Investigation of the gastric mucosa spirochetes of the human," *Gastroenterology*, vol. 27, no. 2, pp. 218–220, 1954.

[19] J. R. Warren and B. Marshall, "Unidentified curved bacilli on gastric epithelium in active chronic gastritis," *The lancet*, vol. 321, no. 8336, pp. 1273–1275, 1983.

[20] A. S. Freedberg and L. E. Barron, "The presence of spirochetes in human gastric mucosa," *The American Journal of Digestive Diseases*, vol. 7, no. 10, pp. 443–445, 1940.

[21] H. A. Alturkistani, F. M. Tashkandi, and Z. M. Mohammedsaleh, "Histological stains: a literature review and case study," *Global journal of health science*, vol. 8, no. 3, p. 72, 2016.

[22] E. Loken and A. Gelman, "Measurement error and the replication crisis," *Science*, vol. 355, no. 6325, pp. 584–585, 2017.

[23] M. Ahsanullah, B. Kibria, and M. Shakil, "Normal distribution," in *Normal and Student st Distributions and Their Applications*, pp. 7–50, Springer, 2014.

[24] M. Weisberg, "Robustness analysis," *Philosophy of science*, vol. 73, no. 5, pp. 730–742, 2006.

[25] M. Thicke, "Evaluating formal models of science," *Journal for General Philosophy of Science*, vol. 51, no. 2, pp. 315–335, 2020.

[26] T. C. Schelling, "Models of segregation," *The American economic review*, vol. 59, no. 2, pp. 488–493, 1969.

[27] C. Mayo-Wilson and K. J. Zollman, "The computational philosophy: simulation as a core philosophical method," *Synthese*, vol. 199, no. 1, pp. 3647–3673, 2021.

[28] A. Adiga, D. Dubhashi, B. Lewis, M. Marathe, S. Venkatramanan, and A. Vullikanti, "Mathematical models for covid-19 pandemic: a comparative analysis," *Journal of the Indian Institute of Science*, vol. 100, no. 4, pp. 793–807, 2020.

[29] I. Kant, *Moral law: Groundwork of the metaphysics of morals*. Routledge, 2013.