Johanna Thoma

The epistemic division of labor revisited

Article (Accepted version)
(Refereed)

The Epistemic Division of Labor Revisited[1]

Johanna Thoma[2]

Abstract:

Scientists differ in the ways they approach their work. Some are happy to follow in the footsteps of others, and continue with work that has proven fruitful in the past. Others like to explore novel approaches. It is tempting to think that herein lies an epistemic division of labor conducive to overall scientific progress: The latter, explorer-type scientists, point the way to fruitful areas of research, and the former, extractor-type scientists, more fully explore those areas. And indeed, it has now long been acknowledged that the social structure of science can play an important epistemic role. Still, philosophers of science have so far failed to produce a model that demonstrates the epistemic benefits of such division of labor. In particular, Weisberg and Muldoon's (2009) attempt, while introducing an important new type of model, suggests that it would be best if all scientists were explorer-types. I argue that this is due to implausible modeling choices, and present an alternative agent-based 'epistemic landscape' model which succeeds at showing the alleged epistemic rewards from division of labor, with one restriction. Division of labor is only beneficial when scientists are not too inflexible in their choice of new research topic, and too ignorant of work that is different from their own. In fact, my model suggests that the more flexible and informed scientists are, the more beneficial is division of labor.

[2] University of Toronto, Department of Philosophy, 170 St. George Street, Toronto, ON M5R 2M8, Canada; johanna.thoma@mail.utoronto.ca

1. Introduction

Consider this stylized account of the Keynesian Revolution, as it is accepted and taught by many economists today.[3] There seems to be a consensus that in the decades before the Keynesian Revolution, there was little progress in English-language economics. The dominant school was Neoclassical Economics, with Alfred Marshall as its figure-head. Neoclassical Economics, in Marshall's own eyes, was simply an extension and elaboration of Classical Economics. Cornerstone were supply and demand diagrams that were used to study partial equilibria, that is, equilibria in individual industries of the economy. These models, however, assumed that markets always clear, so that they could not explain the sustained unemployment of the 1930s. What was needed, it seems, was somebody to try a radically new approach. That someone turned out to be John Maynard Keynes, with his 1936 *General Theory of Employment, Interest, and Money*. But the book was initially received with skepticism, and it lacked formal models and concrete policy advice. The contributions of young economists such as John Hicks, Paul Samuelson and Alvin Hansen were needed to make the Keynesian Revolution the success that it was. So we might say that two things were crucial for progress in 1930s economics: An adventurous economist with novel ideas, as well as followers, who were willing to extract all the important results that can be established when those novel ideas are employed.

The motivation for the model I am going to present in this paper is the observation that, as in this story of the Keynesian Revolution, both explorer-type and extractor-type scientists have played important roles in the history of science, both when we think of

---

[3] See Mankiw (2006) for a typical rendering by a contemporary economist. For a more subtle analysis, see Blaug (1997).

discipline-wide shifts in research program, and different roles within research teams. It is tempting to think that herein lies a division of labor that benefits the epistemic community at large: explorer-type scientists point the way to fruitful new areas of research, and extractor-type scientists extract all the important results. And then scientific progress is faster than it would have been under different research regimes.

Weisberg and Muldoon (2009) set out to demonstrate that there are such epistemic benefits to this type of division of labor with an agent-based 'epistemic landscape' model. However, a closer reading of their paper shows that their model fell short of this task. While the model suggests that the presence of explorer-type scientists makes extractor-type scientists more productive, it also seems to show that it would be better still if everybody was an explorer-type. I present an alternative agent-based epistemic landscape model which succeeds at showing the alleged epistemic rewards from division of labor under many circumstances. Moreover, the model I present is not an ad hoc modification of Weisberg and Muldoon's model: I argue that their failure to demonstrate the benefits of division of labor is due to implausible modeling choices, which my model aims to improve on. In addition, my model allows us to further study the conditions under which division of labor is beneficial. In particular, it suggests that division of labor is only beneficial when scientists are not too inflexible in their choice of new research topic, and too ignorant of work that is different from their own. In fact, the more flexible and informed scientists are, the more beneficial is division of labor.

The paper proceeds as follows: Section 2 introduces epistemic landscape models, and in particular Weisberg and Muldoon's. It presents their results, and conjectures as to why they failed to produce the result they aimed for. Section 3 explains my alternative, and as I will

argue, more credible model. Section 4 presents and discusses the results from a number of simulations I ran on it, and Section 5 concludes.

2. Past Models of the Epistemic Division of Labor

It has been acknowledged for some time that division of labor between scientists may be important for the epistemic progress of a community. Division of labor always involves some kind of diversity in the type of work individual agents do, which helps a group to better achieve a task. In the case of science, this task is an epistemic one. Economists have long been interested in how dividing labor can help a group become more productive. And so philosophers have naturally looked to economic models and applied them to science in order to study the benefits of division of labor, and the ways of sustaining it. It is important to note that there are different kinds of division of labor that may play a role in science. What earlier models focused on was *diversity in the type of research that is conducted*. Kitcher (1990), as well as Strevens (2003) produced rational choice models in which scientists choose between two competing research programs. For the community it is best if there are scientists working on both, but scientists need to be swayed by non-epistemic rewards in order to not all choose the more promising project.[4] Zollman (2010) offered a Bayesian model of consensus formation concerning the efficacy of a treatment, where diversity in treatment method for some time is needed to ensure consensus eventually forms on the right treatment. Hong and Page (2004) use an agent-based model to show that a population of scientists that uses a diversity of problem-solving strategies is better at exploring some search space.

---

[4] See Muldoon and Weisberg (2011) for a critique of this approach.

The type of division of labor that Weisberg and Muldoon, as well as my model focus on is different: It concerns not diversity in research approach, but *diversity in the way scientists go about choosing their research approach*. Scientists differ not only in the type of research they do, but also in how adventurous they are when choosing what kind of research to do. And this, too, could amount to an epistemically beneficial division of labor.[5] To highlight that this is a different kind of division of labor, I will use the phrase 'diversity of research strategy' for this second type of division of labor throughout. Agent-based epistemic landscape models, as Weisberg and Muldoon introduce them, lend themselves very well to modeling diversity of research strategy. By offering a spatial representation of similarity of research approach, they allow us to represent explorer-type and extractor-type behavior as different ways of moving on an epistemic landscape. Epistemic landscape models also provide a tractable representation of a fine-grained research field, in which scientists can choose between a large number of research approaches. Given such a research field, the agent-based approach allows us to study the dynamics of the different strategies concerning choice of research project over time.[6]

2.1. Epistemic Landscape Models

---

[5] De Langhe (2014) aims to address both types of division of labor in one model. Unlike Weisberg and Muldoon and I, he does not use an epistemic landscape model. In his agent-based model, agents weigh up the benefits of exploiting existing theories, and exploring new ones. He argues that an adaptive strategy, where agents 'explore' or 'exploit' depending on the relative costs and benefits is in fact ideal for maintaining the right level of diversity in the theories that the community is working on.

[6] Muldoon (2013) provides a more detailed overview of different models of the epistemic division of labor, as well as the advantages of epistemic landscape models.

The basic idea behind epistemic landscape models is the following: An epistemic landscape represents a research field. Scientist-agents are modeled as making discoveries within that research field by moving around on the epistemic landscape. Different points in the epistemic landscape represent scientific *approaches* with an associated epistemic *significance.* An approach is characterized by a methodology (regarding data collection as well as data analysis), a research question and a set of background beliefs. Epistemic landscapes represent similarity of research approach as spatial proximity. Weisberg and Muldoon's epistemic landscape is two-dimensional and discrete, so that different research approaches are represented as discrete patches on a plane, arranged on a grid of 101 x 101 patches (p.234). Even though research approaches are assumed to be discrete, we can interpret them to be very fine-grained: For instance, using the same methodology and background assumptions as somebody else in order to treat a slightly different research question could count as a different, but nearby approach.

The epistemic significance associated with each approach is a numerical value which represents, roughly, the amount of scientifically or socially important results that can be obtained using a particular approach.[7] It represents what the epistemic community cares about. Weisberg and Muldoon's model assumes that epistemic significance is not distributed randomly on the landscape, but in two Gaussian-shaped hills with single peaks, where one of the peaks is higher than the other (p.234). This makes sense if we assume that approaches that

---

[7] See Kitcher (1993) on the concept of epistemic significance. Weisberg and Muldoon stress that epistemic landscape models do not rely on any particular interpretation of epistemic significance, but do assume that we all share the same conception of epistemic significance, and correctly identify it (p.229). This is because all agents make their behavior depend on one single value of epistemic significance for each approach, which is the same for all. Of course this is an idealization, which could be relaxed in further studies.

are similar to significant approaches are also likely to be similarly significant, and that there are several areas of epistemic significance within one research field. Large parts of the epistemic landscape are assumed to be entirely insignificant. Figure 1 shows an epistemic landscape with patches colored according to significance.
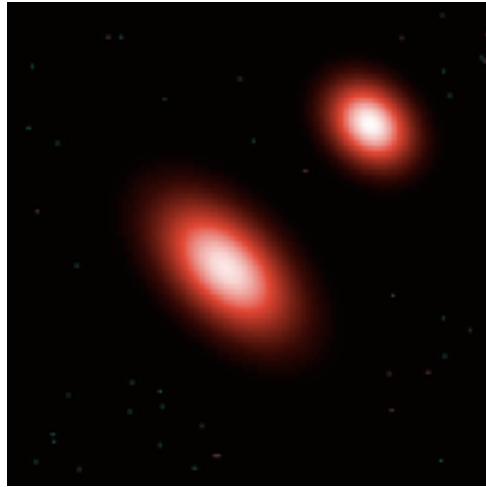
Figure 1: An epistemic landscape

Scientific progress is now modeled as occurring when scientist-agents, who are initially randomly distributed on the insignificant areas of the landscape, move around the landscape making discoveries. When scientists 'visit' a patch, they use the approach of that patch to find out its significance. Weisberg and Muldoon assume that all agents successfully determine the significance of an approach when they use it (p.233). Scientists can move around the landscape according to rules. By specifying different rules, we can study how well different types of scientists do at exploring the epistemic landscape, and how they interact, in the hope of thereby learning something about scientific progress in the real world.

2.2. Followers and Mavericks

Weisberg and Muldoon implement three rules for scientists: a Control rule, a Follower rule, and a Maverick rule. Each type of scientist-agent responds to information about their own past discoveries, and can only ever move forward one patch at a time. For instance, Controls move straight ahead unless their current approach is less significant than their last, in which case they go back and change direction (pp. 231-232). They thus behave like 'hill-climbers'. But Mavericks and Followers also respond to information about what discoveries have been made by others in their vicinity. These two rules aim to capture the explorer-type and extractor-type behavior we described earlier. Mavericks are explorative in that they avoid approaches that have been used previously. Unless they start going 'downhill', in which case they go back and change direction, they move to a random *unvisited* patch in their neighborhood (p.243). Followers, on the other hand, like to move to approaches that have proven fruitful in the past. They move to the *previously visited* patch with the highest significance in their neighborhood. If there are none, they move to a random neighboring patch (pp. 239-240).

Weisberg and Muldoon run a number of computer simulations to study how well differently composed populations of scientist-agents do at finding the significant areas of the landscape. They use three measures of success: First, they look at how long it takes the different populations to find both peaks of the landscape (p.234). Second, they record what proportion of approaches with positive significance have been found by different populations after different time periods. They call this *epistemic progress* (p.237). Third, they look at what proportion of the total number of approaches have been discovered by the different populations after different time periods. They call this *total progress* (p.248).

2.3. Weisberg and Muldoon's Results

Total Progress of 400 Scientists
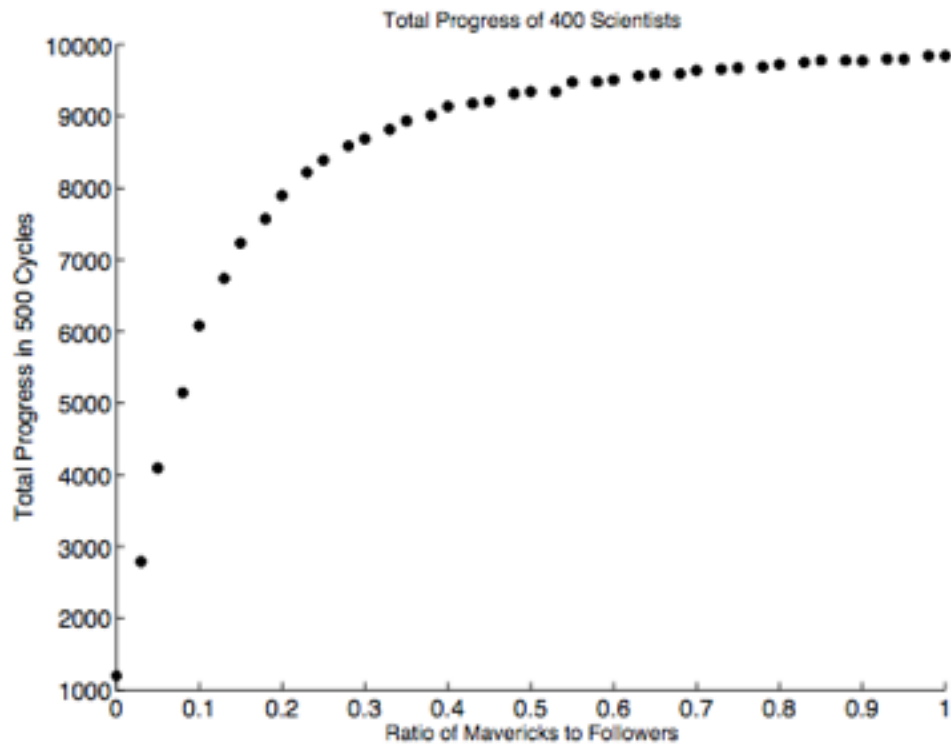
One major result from the simulations Weisberg and Muldoon ran is that homogenous populations of Mavericks do better than homogenous populations of Controls (pp. 244-245). So agents do better when they take into account in some way information about what work others have previously done. This already shows the benefits of division of labor in one sense of the word. However, we would also like evidence that *diversity of research strategy* is epistemically beneficial. And to get such evidence, it seems that we need to show that mixed populations of Followers and Mavericks do better by the various criteria of success than homogenous groups. This is also what Weisberg and Muldoon seem to suggest they show in their paper. In the abstract they write,

> "[W]e show that in mixed populations, mavericks stimulate followers to greater levels
> of epistemic production, making polymorphic populations of mavericks and followers

ideal in many research domains."

However, if polymorphic populations are indeed ideal, their simulations do not show this. In fact, according to their own measures of success, homogenous populations of Mavericks outperform all mixed populations. Figure 2 shows the total progress made by differently

Figure 2: Total progress of different populations of 400 scientists after 500 rounds

composed populations of 400 scientist-agents after 500 periods. What the data seem to show is that the more Mavericks there are in a population, the better.[8]

Weisberg and Muldoon point out that the Maverick strategy is likely to be more costly than the Follower strategy. Taking into account those costs, they claim, a mix of Followers and Mavericks will turn out to be optimal (pp. 250-251). However, they do not present a formal

---

[8] Weisberg and Muldoon do not present data on the *epistemic* progress made by differently composed populations where the number of agents is held fixed. What they do present is data on what happens to epistemic progress when Mavericks are added to populations of Followers (pp.246-247). Epistemic progress is significantly increased by adding Mavericks, and Mavericks also have an indirect positive effect on the performance of Followers within mixed groups. However, when we compare this to the data on homogenous groups of Mavericks presented earlier in the paper (p. 245), we see that these mixed populations do not do as well in terms of epistemic progress as homogenous groups of Mavericks of a similar size.

model that incorporates these costs. And looking at the difference in productivity between Followers and Mavericks, doubts should arise about their claims. All homogenous groups of Mavericks eventually find virtually all significant approaches in the landscape. For instance, within 200 periods, 400 Mavericks find more than 90% of the significant approaches (p. 245). In contrast, homogenous groups of Followers all eventually stagnate at a low value - 17% in the case of 400 Followers (pp. 240-41). Note that in figure 2, homogenous populations of Followers discover only slightly more than 1,000 approaches, which means that each scientist on average only discovers around 3 new approaches within 500 periods, suggesting that they either hardly move, or spend the majority of their time on patches that have previously been explored. Even if it is granted that Followers become more productive when they are stimulated by Mavericks, it is simply not obvious that it would pay off for an epistemic community to include Followers in the scientific community, even if their research strategy is significantly cheaper.

2.4. Problems with Weisberg and Muldoon's Model

We just saw that Weisberg and Muldoon's model does not clearly demonstrate the benefits of diversity of research strategy, and rather suggests that it would be best for all scientist-agents to be Mavericks. This also raises a question at the micro-level: Assuming that scientists are motivated in large part by the desire to make significant discoveries, if Mavericks are so much better at making significant discoveries, why would anybody want to be a Follower? The Follower strategy turns out to be extremely unattractive in this model. But that does not match what we observe in scientific practice: A large proportion of scientists does seem to be engaged in what we described above as extractor-type research, that is, research that is

somehow based on or similar to work that has previously proven fruitful. This raises doubts whether Weisberg and Muldoon's model accurately captures extractor-type and explorer-type behavior in actual scientific practice.

Weisberg and Muldoon's specification of the Follower rule in particular seems to me to be implausible given the other modeling assumptions they make. On their Follower rule, scientists end up duplicating the work of others much of the time: They are characterized by a tendency to move to the most successful previously explored approaches in their neighborhood. But mere duplication has no added benefit in the setup of the model. As we have seen, Weisberg and Muldoon's measures of success only count the first time an approach is used.[9] Given this modeling assumption, any strategy which has scientists merely duplicate others' work much of the time seems implausible as a representation of actual scientific practice, for two main reasons: First, research results are mostly freely available to other scientists, so we can assume that researchers already have access at least to the research results from approaches that are similar enough to their own for them to understand and easily find them. Second, if this is so, it is hard to see why anybody could be motivated to simply duplicate the work of others: There is no epistemic benefit from doing so. After Hicks read the *General Theory*, he followed Keynes' lead, but he did not write the *General Theory* once over again.

---

[9] Relaxing this assumption may be another way in which one could make Weisberg and Muldoon's Follower strategy more plausible: If the first person to use an approach didn't uncover all significant results that can be obtained with that approach, there would again be a point to Follower behavior. And in fact, there are a number of rewards to duplication, not least that it helps to detect error. But unless the model makes these benefits explicit, the Follower strategy has scientists engage in 'pointless' behavior much of the time. Rather than making these benefits explicit, I below choose the option of changing the strategy to avoid duplication in order to keep as much as possible of the original framework fixed and for ease of presentation.

There is another aspect of Weisberg and Muldoon's specification of the Follower and Maverick strategies which seems implausible, namely that all agents only move locally. Agents only respond to previous discoveries in their immediate neighborhood, and they can only ever move one patch ahead. One way to justify this assumption would be by claiming that movement at a larger range is costly, and that information about approaches different from one's own is hard to get by. It is surely the case that, even though much information is now freely accessible, acquiring detailed knowledge of areas of research very different from one's own is hard. And it is also hard to acquire the new research tools needed to work on a very different approach. As Muldoon and Weisberg (2011) highlight, one of the advantages of using an agent-based epistemic landscape model is that we can represent agents as being restricted to a specific research area in this way. But for strict local movement to be a plausible description of actual scientific progress, it would need to be the case that agents are restricted to the approaches that are immediately adjacent to their own. And on the most plausible interpretation of the model, this would amount to an extreme level of short-sightedness and inflexibility among scientists.

Whether local movement is plausible depends to some extent on how fine-grained we interpret research approaches to be. If they are coarse-grained, it is more plausible than if the distinctions between neighboring approaches are very small. But given the way the model is implemented, we need to in fact think of the epistemic landscape as very fine-grained. We said that a single scientific agent can discover the entire significance of one approach in one round. If approaches were coarsely grained, this would be unrealistic: It would then have been better to allow for a patch to be visited multiple times, or for single agents to stay on a patch for longer. Of course, we could think of each round as lasting a long time. But what Weisberg and

Muldoon presumably aimed to capture is a more detailed look at the dynamics of research groups, that tracks smaller changes in research approach. Imagine we interpret each round as lasting about 2 months. In that case, the 200 rounds that 400 Mavericks need in order to discover 90% of the significant approaches would come to about 33 years of research, which seems like a plausible lifetime for a specialized field of research. Given that interpretation, local movement would imply that scientists are not aware of research that they would be able to conduct themselves within 4 months of research. For instance, in our earlier example, this would mean that somebody like Hicks could become aware of an early paper of Keynes' on a topic that he himself happens to be working on, and remain unaware that Keynes has already written the *General Theory*. Similarly, scientists could not now learn the research tools for using an approach that they will be able to use after just two month of research on something else. All in all, I think that given the fine-grained interpretation of research approaches that is implicit in the model, it is more plausible to give scientist-agents a larger range of movement, while still acknowledging some limitations to awareness and research flexibility.

3. An Alternative Epistemic Landscape Model

We just identified two potential flaws in Weisberg and Muldoon's model that may be responsible for their failure to demonstrate the epistemic benefits of diversity of research strategy. Firstly, their specification of the Follower strategy is implausible because it has Followers duplicate the work of others much of the time. And secondly, the assumption of local movement seems too restrictive. I hence constructed an alternative epistemic landscape model that differs from Weisberg and Muldoon's in two major ways: Firstly, it allows for a variable range of movement of the scientist-agents, where the range of movement can be

understood as the degree to which scientists are *flexible* in their choice of new research approach, and *informed* about work that is different from their own. Secondly, it uses strategies according to which all scientists avoid mere duplication of previous work. Given these changes, I indeed found diversity of research strategy to be beneficial, with one restriction. Diversity of research strategy is only beneficial when the range of movement is not too small, that is, when scientists are not too inflexible in their choice of new research topic, and too ignorant of work that is different from their own. In fact, my model suggests that the more flexible and informed scientists are, the more beneficial is diversity of research strategy.

We can still meaningfully distinguish between extractor-type and explorer-type scientists even when all avoid duplication, in the following way. My model describes explorer-types as scientists who like to follow approaches that are very different from those of others, while extractor-types like to do work that is very similar to but not the same as that done by others. Epistemic landscape models lend themselves well to programming these kinds of strategies, due to their implicit similarity measure. I call the two respective strategies in the model the *Explorer* and the *Extractor* strategy. Given these two types of strategies, I investigated under what conditions mixed populations of Extractors and Explorers are better at making scientific discoveries than homogenous populations of Extractors and Explorers respectively. Keeping Weisberg and Muldoon's assumptions fixed unless otherwise stated, I used the programming software NetLogo (Wilensky 1999) to implement the model and run simulations on it.

The following is a more detailed description of the Explorer rule when movement is local. First suppose Explorers are in unchartered territory, so that no approaches in their neighborhood have previously been explored. In that case, Explorers behave like hill-climbers:

They go straight ahead unless they start going downhill, in which case they go back and change direction. Now suppose that some of the approaches in their neighborhood have previously been explored. In this case, they go to the unvisited patch at the *greatest minimum distance* to the patches that have previously been *visited by others* (as illustrated in figure 3). If there are no unvisited patches, they go to a random neighboring patch.[10]
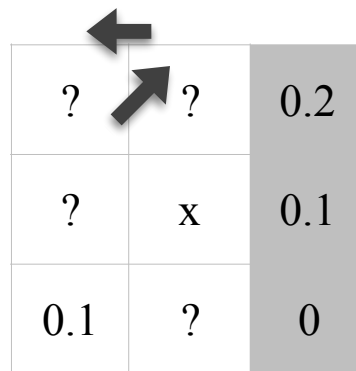

Figure 3: Movement of an Explorer (x): The grey fields have been previously visited by other scientists.

The Explorer strategy describes the behavior of a scientist who does care about making significant discoveries - after all Explorers behave like hill climbers when they are in unchartered territory. But the scientists the rule describes also like to move away from research that other scientists are doing. This is captured by the Explorers' behavior when they are in the

---

[10] If they moved to the best previous discovery rather than a random one, they could easily get trapped in local maxima, which can't be what scientists concerned with making significant discoveries want. When they are in an area that has been explored before, they should aim to find an unexplored area. I model this by assuming that they engage in a random walk until they find an unvisited patch.

vicinity of previously discovered patches, and by their tendency to go straight ahead once they have set their direction away from those patches.

The Extractor Rule, on the other hand, describes the behavior of a scientist who likes to do work that is similar, but not identical to work that was successful in the past. Suppose again that the agent is in unchartered territory. The Extractor then moves to a random unvisited patch in the neighborhood (if she has visited them all herself, she goes to a random neighbor). Now suppose that approaches in her neighborhood have been visited by others. As figure 4 illustrates, the Extractor then moves to the unvisited patch closest to the best patch discovered by another scientist (like the Explorer, if all patches have been visited, she goes to a random neighbor). Since more significant approaches are likely to be in the vicinity of other significant approaches, this also expresses a desire to make significant discoveries.
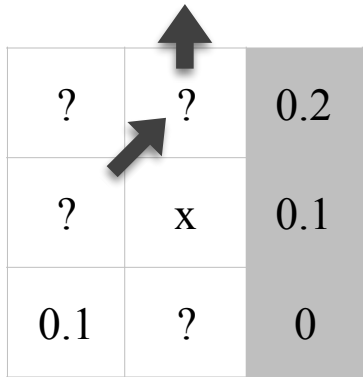

Figure 4: Movement of an Extractor (x): The grey fields have been previously visited by other scientists.

Both strategies can easily be extended to non-local movement by exchanging any reference to the neighborhood in the above descriptions for a specified radius. So Explorers

move to the unvisited patch *within radius r* that is at the greatest minimum distance to the patches others previously visited *within radius r.* And Extractors move to the patch *within radius r* that is closest to somebody else's best previous discovery *within radius r*. Note that this alters not only the agents' range of movement, but also their range of 'perception', that is the range within which they are aware of other scientists' research.[11]

4. Results

Using the strategies just described, I ran a number of simulations recording two sets of data. Firstly, I recorded what proportion of the total significance contained in the research field each population discovered after various periods of time. This is the measure of success we assume the epistemic community cares about. Note that this is different from Weisberg and Muldoon's criteria of success. However, it nicely integrates two concerns, since it is sensitive both to the number of significant patches discovered, and the relative significance of the approaches that are discovered. Secondly, I recorded data on how much significance each type of scientist discovers, on average, each period. This allows us to determine the relative productivity of each type of scientist, which matters for how attractive the strategies are for scientists. Simulations were run for various group sizes, group compositions, and ranges of movement, with 50 repetitions each.

4.1. Local Movement

---

[11] Of course, the two ranges could come apart. Exploring what happens in these cases would be an interesting area for further investigation.

The results for the case of local movement were as follows: All groups of size 40 and larger found approximately the entire epistemic significance of the research field within 500 rounds or less. This is reassuring, since it means that all strategies are reasonably efficient, in contrast to Weisberg and Muldoon's Follower strategy. Still, epistemic communities do not only care about eventually finding the entire significance, which all groups will at some point, but about making progress fast. And different groups of scientists differ greatly in that respect. For instance, unsurprisingly, larger groups do better than smaller ones. Furthermore, homogenous groups of Explorers do better than homogenous groups of Extractors, especially early on in the simulations. Lastly, at least until late in the research cycle, homogenous groups of Explorers also do better than mixed groups. Figure 5 illustrates this in more detail, in the case of a group with a 50/50 mix of Explorers and Extractors.

Figure 5: Proportion of significance discovered by groups of 40 scientists
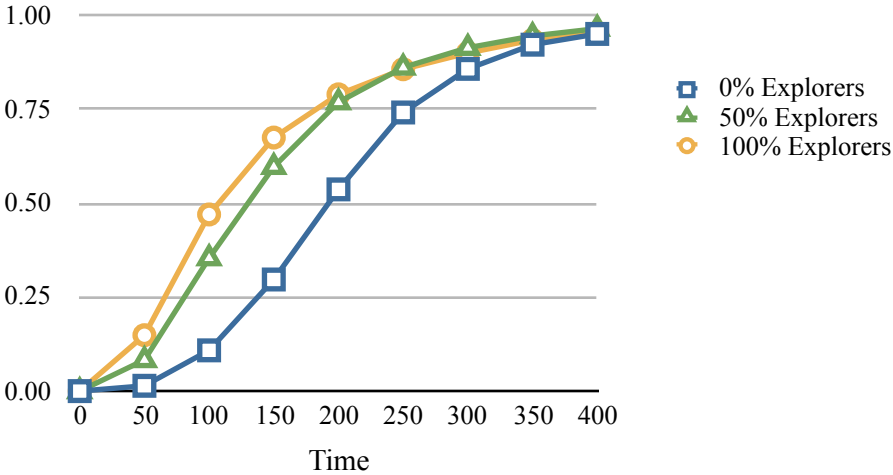
The mixed group does better by a very small margin after 300 rounds. But at this point, 90% of the total significance has already been discovered. Homogenous groups of Explorers

have a large advantage earlier on in the research cycle. For instance, after 50 rounds, the homogenous groups of Explorers have discovered 15%, while the mixed group only discovered 8% of the total significance. This seems to outweigh the small advantages mixed groups have later on, especially if we assume that the epistemic community cares in a special way about making early discoveries fast. Looking at the average significance of moves by Explorers and Extractors, it turns out that Explorers also outperform Extractors within the mixed groups.

4.2. Medium Range Movement

The upshot from what I just presented seems to be that when movement is local, diversity of research strategy is not beneficial. But I have also argued above that local movement is unrealistic because of the extreme short-sightedness and ignorance it implies. The more realistic case is one where agents can move at a larger range. And in fact, things change when we allow for a larger range of movement. From a range of movement of 3 upwards, mixed groups start to do best, followed by homogenous groups of Extractors, with homogenous groups of Explorers doing worst. In fact, while the range of movement is still restricted,[12] the larger the range of movement, the faster the overall progress and the more beneficial diversity of research strategy. Figure 6 presents the results for a range of movement of 10 in more detail. Clearly, mixed populations of Explorers and Extractors perform better than homogenous

---

[12] For computational reasons, I only ran simulations up to a range of movement of 10.

populations at various time points.[13] This suggests that division of labor between Explorers and Extractors is indeed beneficial when movement is no longer local.
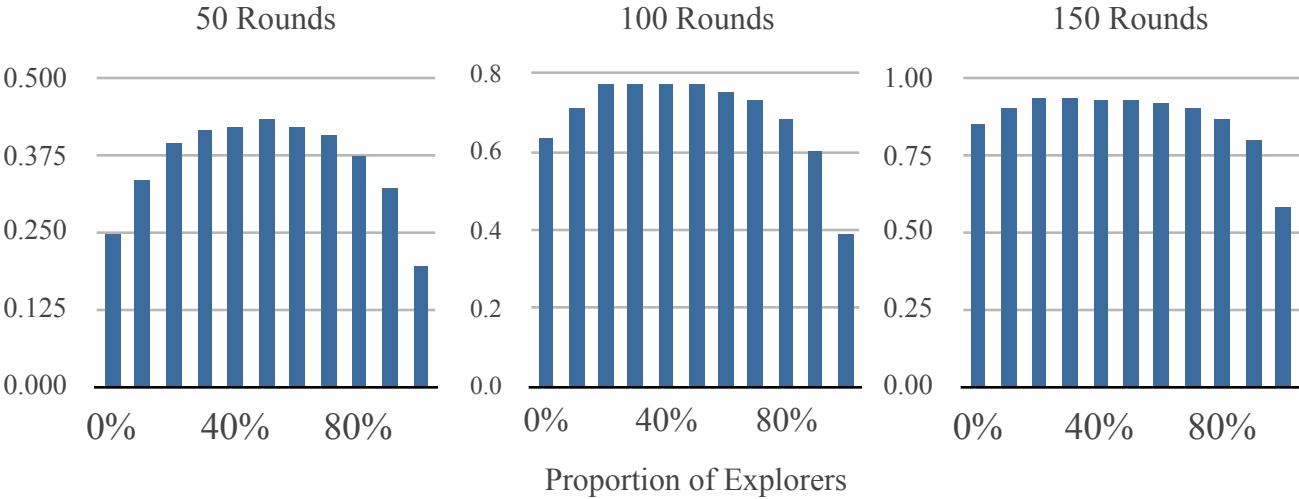

Figure 6: Proportion of significance discovered by populations of different composition after 50, 100 and 150 rounds

While the advantage of mixed groups over homogenous groups of Extractors is not as large as their advantage over Explorers in terms of the *average* proportion of significance discovered after various time points, part of the advantage of adding Explorers to groups of Extractors is that the standard deviation is greatly reduced. For instance, after 150 time steps, the standard deviation for homogenous groups of Extractors is 12.5 percentage points, while that for mixed groups with 40% Explorers is only 3.1 percentage points. This reduced

---

[13] Two-sample t-tests (assuming unequal variances) show that the differences in average significance discovered between a 50/50 mix and each of the homogenous populations at each time point are all significant with a p-value below 0.0001.

variability should also make mixed groups more attractive, since there is less risk involved for the epistemic community.

What could explain the difference between the case of local movement and the case of medium range movement? Figure 7 depicts the paths taken by Explorers and Extractors in a typical simulation when the range of movement is 1 (left) and 10 (right).
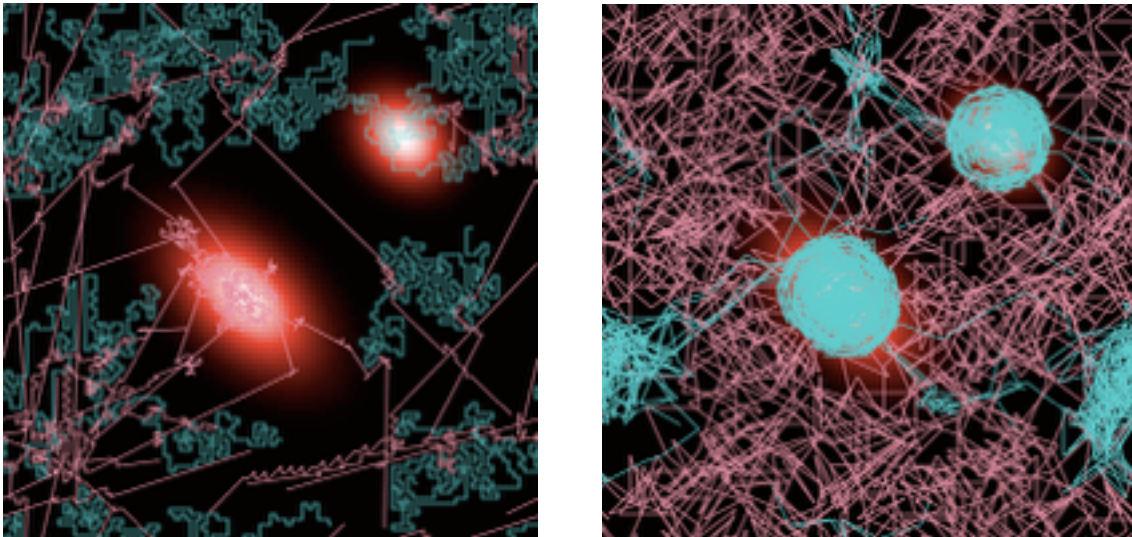
Figure 7: Local movement (left) vs medium range (range 10) movement (right). The paths of Extractors are colored blue.

In the case of local movement, it appears that Extractors mostly do not manage to follow the paths of the most successful Explorers, while they do so in the case of medium range movement. Two explanations suggest themselves for this: Firstly, since their range of 'perception' is limited, they cannot 'see' which paths will lead to the most significant areas. Secondly, their choice of who to follow is very restricted - they have to be very lucky for a

successful Explorer to pass through a patch right beside them. This not only makes sense of what happens in the model, but sounds like a plausible story about why diversity of research strategy might not be beneficial when real world scientists are too uninformed and short-sighted: Visionary explorer-types are no good for guiding the way when extractor-types do not become aware of their work, or do not recognize it as visionary.[14]

4.3. Global Movement

Agents who move at a range of 10 are still restricted in how informed they are about others' research and in their flexibility in adopting new research approaches. We may be interested in what happens when agents can move globally, that is, in the entire research field. To study global movement, I had to slightly adapt the Explorer rule for computational reasons: After trying for a while to find an approach that is far away from previously discovered approaches, they move to a random unvisited approach somewhere in the research field. Some pilot simulations with the original rule suggested that there is no significant difference between this rule and the original one. What happens in the global movement case is the following:

---

[14] There is an alternative explanation of the difference between local and medium range movement according to which it is just an artifact of the 2-dimensional nature of the model. We might think that Extractors do poorly in the local case because they get 'fenced in' by the paths left by Explorers. Explorers leave many straight paths that Extractors cannot 'jump over' when movement is local. But this cannot be the whole story. Diversity of research strategy is still not unambiguously beneficial at a range of movement of 2, which allows for 'jumping over fences', and we also saw that the benefits of diversity of research strategy become larger as the range of movement becomes larger. The story we told above makes sense of this pattern, too. Furthermore, I ran simulations where Extractors are allowed to temporarily jump over paths when it looks like they are being 'fenced in', and diversity of research strategy still did not turn out to be beneficial.

Diversity of research strategy is still beneficial, but overall progress is slower again compared to the medium range movement case. The explanation for this seems to be that the lower of the two hills gets neglected by Extractors as soon as the higher one is found.[15]

These results suggest that, if the epistemic community cares about exploring all areas of significance, there may be an advantage to scientists being slightly inflexible and/or uninformed, and so moving at a medium range. Medium range movement also seems to be the more credible representation of actual scientific practice. Due to the costs of changing one's research approach, scientists are not usually flexible enough to adopt just any new approach within their research field. To take again the example of the Keynesian Revolution, Keynes' followers did not move away completely from Neoclassical Economics, and certainly employed some of the neoclassical tools they were used to. They were later responsible for what is now called the 'Neoclassical Synthesis'. Similarly, it would be very demanding to stay on top of all research that happens within one's field of research, especially when that field is large. Hence a medium range of movement seems most realistic.

_____

[15] As an anonymous referee pointed out, this is very similar to the problem that Strevens (2003) and Kitcher (1990) aim to address. They are interested in how scientists could be induced to work on the less promising of two research projects. Strevens studies how different reward structures could help to achieve an efficient allocation of researchers. The Explorer strategy in my model is not responsive to the types of incentives he studies, and thus too crude to handle the problem in this way. However, my model suggests that the problem could also be avoided if agents move at a medium range.

4.4. Average Performance of Explorers and Extractors

Looking at the data on the relative performance of Explorers and Extractors within the mixed groups, we can see that, apart from a very short period in the beginning, Extractors do better than Explorers - and this advantage is bigger the larger the range of movement. At range 10, Extractors are more than 4 times as productive as Explorers. This raises a question similar to one that we raised for Weisberg and Muldoon. In their model, it was unclear why anybody would choose to be a Follower, given their lack of productivity. In our case, the question is why anybody would choose to be an Explorer. While, as we have seen, it is certainly beneficial from an epistemic community's point of view to include explorer-type scientists in the scientific community, the strategy is not very attractive: It is both riskier, and has a much lower average pay-off in terms of epistemic significance.

In order to give a full response to this challenge we would have to include the scientists' choice of strategy explicitly in the model, which we have not done. My model simply assumed that a fixed number of agents follow a particular research strategy. This simplification seems appropriate given our main goal of showing that it is conducive to scientific progress to have a mix of explorer-type and extractor-type scientists. Still, if one strategy turns out to look so unattractive that it seems unlikely that any scientist would ever choose it, this is a problem for the model. We want to say that the model gives evidence that division of labor between extractor-type and explorer-type scientists in actual scientific practice is beneficial. And some scientists do choose to be explorers, apparently for good reasons. If the model seems incompatible with this, it may not be a good representation of scientific practice after all. So we should at least have some response to this charge, even if the choice of strategy is not part of our model.

I think in contrast to Weisberg and Muldoon's case, a plausible response can be given to the challenge. And that is that there are a number of social incentives that make the Explorer strategy more attractive from the individual point of view. We mentioned at the outset that Kitcher (1990) and Strevens (2003) constructed models that suggested that non-epistemic rewards are needed to maintain an epistemically beneficial diversity of research approach. Perhaps something similar holds here, too. Given that the average epistemic significance discovered by Explorers is so much lower, the strategy will only be attractive to somebody who cares about something other than epistemic significance. And so individual scientists, if they should choose to become an explorer-type, cannot only care about what the epistemic community at large cares about.

In fact, many scientists are also interested in monetary and social rewards, like prestige.[16] And, arguably, these kinds of rewards are given disproportionately to scientists who do the kind of work that Explorers do. For instance, the novelty of the proposed research is an important desideratum for most grant-giving agencies. In the EU, the ERC has even made it its main mission to support high-risk 'frontier-research' (note again the geographical analogy). Furthermore, fame is often associated with having come up with something very new.

My model is consistent with the claim that these social and financial rewards could have an important function: They could help to maintain an epistemically beneficial diversity of research strategy by making sure that the Explorer strategy is attractive enough for some scientists to choose it. This is not to suggest that all revolutionaries, such as Keynes, were motivated by fame and money. Even in the absence of social and financial rewards, there may just be something inherently more attractive about the Explorer strategy compared to the

---

[16] see Fox (1983) and Levin and Stephan (1991) for empirical evidence for this claim.

Extractor strategy, which can draw some scientists to it. There is certainly a thrill in doing something new.

5. Conclusion

The agent-based epistemic landscape model I have presented here appears to show the following: Division of labor between explorer-type and extractor-type scientists is beneficial whenever scientists are not too inflexible and uninformed about other scientists' research. When scientists are too inflexible and uninformed, extractor-types may not manage to follow explorer-types to more fruitful areas of research. On the other hand, when scientists are perfectly informed and flexible, there is a danger that some potentially fruitful areas of research are ignored.

The model I have presented not only supports an intuitive result that Weisberg and Muldoon's could not. It is also more credible than Weisberg and Muldoon's in two major ways: Firstly, it is not restricted to local movement, which I have argued is implausible as a representation of scientific practice. And secondly, the Explorer and Extractor strategies are better descriptions of the behavior of scientists than the Maverick and Follower strategies, since both Explorers and Extractors avoid the mere duplication of work others have done. What further speaks in favor of the model is that a number of its implications map credibly onto features of actual scientific practice, as evidenced in the course of the paper. For instance, on my model it turns out to be explorer-type behavior which needs special incentives, which seems plausible.

However, the model is of course still highly idealized and could be extended in a number of ways. The results I presented here remain the same when the topology is changed to one with three or four hills, keeping the total amount of significance in the landscape fixed. Still, many other topologies are possible, in particular ones with more than two dimensions. It would also be interesting to look at models that allow for individual differences in flexibility or talent. And lastly, both research costs and non-epistemic rewards could be modeled explicitly in future studies.

References

Blaug, Mark. 1997. *Economic Theory in Retrospect,* 5th Edition. Cambridge: Cambridge University Press.

De Langhe, Rogier, 2014. "A Unified Model of The Division of Cognitive Labor.", *Philosophy of Science* 81 (3): 444-59.

Fox, Mary Frank. 1983. "Publication Productivity Among Scientists: A Critical Review." *Social Studies of Science* 13 (2): 285-305.

Hong, Luo, and Scott E. Page. 2004. "Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers." *Proceedings of the National Academy of Sciences of the United States of America* 101 (46): 16385-16389.

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest and Money.* London: Macmillan (reprinted 2007).

Kitcher, Philip. 1990. "The Division of Cognitive Labor." *The Journal of Philosophy* 87 (1): 5-22.

———. 1993. *The Advancement of Science*. New York, NY: Oxford University Press.

Levin, Sharon G., and Paula E. Stephan. 1991. "Research Productivity Over the Life Cycle: Evidence for Academic Scientists." *The American Economic Review* 81 (1): 114-132.

Mankiw, N. Gregory. 2006. "The Macroeconomist as Scientist and Engineer." *Journal of Economic Perspectives* 20 (4): 29-46.

Muldoon, Ryan. 2013. "Diversity and the Cognitive Division of Labor." *Philosophy Compass* 8 (2): 117-125

Muldoon, Ryan, and Michael Weisberg, 2011. "Robustness and Idealization in Models of Cognitive Labor." *Synthese* 183 (2): 161-174.

Strevens, Michael. 2003. "The Role of the Priority Rule in Science." *Journal of Philosophy* 100 (2): 55–79.

Weisberg, Michael, and Ryan Muldoon. 2009. "Epistemic Landscapes and the Division of Cognitive Labor." *Philosophy of Science* 76 (2): 225–52.

Wilensky, Uri. 1999. NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

Zollman, Kevin. 2010. "The Epistemic Benefit of Transient Diversity." *Erkenntnis* 72 (1): 17-35.