The Epistemic Benefit of Transient Diversity

Kevin J.S. Zollman

University of California, Irvine

kzollman@uci.edu

March 13, 2007

Because of the structure of inductive reasoning, error is an ever-present worry in science. Excluding errors that result from a lack of creativity (where better alternatives remain unconceived), scientific error can be divided into three broad categories. Some occur because of, widely termed, sociological factors. Powerful political parties, cultural biases, or scientific alliances may cloud the judgment (or publication) of scientific results, preventing a superior theory from being accepted. Other errors are the result of scientists' misconduct; data is invented, hidden, or manipulated and as a result others are convinced to accept an inferior theory. Finally, errors may be the result of simple, good faith mistakes on the part of scientists. In this case, everyone did everything right, but as a result of some random occurrence or another, the data was misleading.

These three types of errors can be extremely damaging, especially if an available, superior alternative is dismissed as the result of a scientific error. A discarded theory might never again be revisited; future scientists may remember that this theory was abandoned by their predecessors, and avoid revisiting the decision. A better theory may be lost for long periods of time, perhaps forever. Because of this danger, we hope that the chance of each of these errors persisting is minimized by the behavior of scientists.

One would like to identify mechanisms by which the effect of each type of error is minimized, even if none can be eliminated entirely. This study will examine two ways of reducing the harm resulting from the last type of error, good faith errors. In order to analyze the phenomenon in detail, an interesting historical case of a good faith error in science will be presented. Using this as a paradigm case, we will examine two abstract models of scientific behavior and look at social solutions to the problem of good faith errors.[1]

The analysis of both of these models demonstrates a consistent theme, that a certain amount of diversity provides some benefit to the community. In the first model, we can attain this diversity by limiting the amount of information available to the scientists. This is achieved by arranging them so that they only see a proper subset of the total experiments performed, using different subsets for each agent. This preserves a sort of optimal diversity present in the agents' priors.[2] Such diversity comes at a cost, however, and here that cost is speed. While the groups with access to less information are more reliable, they are much slower at converging to the superior theory.

Similarly in the second model, we can preserve this same diversity by limiting information available to the individuals. Here the difference between the more and less limited groups is even more pronounced. However, this second model also facilitates

[1]These abstract models contribute to the larger project of formal social epistemology (cf. Goldman, 1999). Instead of considering reliability as a property of individuals, this project views it as a property of groups.

[2]The epistemic benefit of diversity is not new with this study. It was recognized by Kitcher (1990; 1993; 2002) and Strevens (2003a; 2003b). The relationship between these models and their's is discussed in Section 4.

the investigation of another avenue for the maintenance of diversity, extreme priors. In this model, we find that if agents have very extreme beliefs, this serves the same purpose as limiting information. In fact, if agents have both extreme priors and only have access to limited information they are less reliable than either feature taken alone.

Overall the moral of these models is that in some cases features of individuals that seem epistemically counterproductive may in fact make communities of individuals more reliable. But, before looking at the model in detail, we will first turn to an actual case of good faith error in science.

1 The Case of Peptic Ulcer Disease

In 2005, Robin Warren and Barry Marshall received the Nobel Prize in Physiology or Medicine for their discovery that peptic ulcer disease (PUD) was primarily caused by a bacteria, *Helicobacter pylori* (*H. pylori*). The hypothesis that peptic ulcers are caused by bacteria did not originate with Warren and Marshall, however, the hypothesis predates their births by over 60 years. Unlike other famous cases of anticipation, it was the subject of scientific tests during that time. To those who have faith in the scientific enterprise, it should come as a surprise that the widespread acceptance of a theory should take so long. If the hypothesis was available and subjected to scientific tests, why was it not widely accepted long before Warren and Marshall? Interestingly, the explanation of this error does not rest with misconduct or pathological science, but rather with simple, and perhaps unavoidable, good faith mistakes.

The first appearance of the bacterial hypothesis occurred in 1875. Two bacteriologists, Bottcher and Letulle advocated that peptic ulcers were caused by an unobserved bacteria. This was supported by observations of bacteria like organisms in glands in the stomach by a another German pathologist. Almost simultaneously, the first suggestions that PUD may be caused by excess acid begin to appear (Kidd and Modlin, 1998).

Before the turn of the century, there are at least four different observations of spirochete organisms (probably members of the Helicobacter species) in stomachs of humans and other mammals.[3]

> Thus, by the turn of the century, experimental results appeared to confirm the hypothesis that a bacterial infection might be, if not an occasional cause, at least an accessory requirement for the development of gastroduodenal ulcers. Thus, although a pathological role for bacteria in the stomach appeared to have been established, the precise role of the spirochete organisms remained to be further evaluated. (Kidd and Modlin, 1998)

During the first half of the 20th century, it appears that both hypotheses are alive and well. Observations of bacteria in the stomach continue, and reports of successful treatment of PUD with antibiotics surface.[4] At the same time the chemical processes of the stomach becomes better understood, and these discoveries begin to provide some evidence that these processes may be involved in PUD. Supporting the hypoacidity theory, antacids are first used to successfully treat PUD in 1915. This success encourages further research into chemical causes (Buckley and O'Morain, 1998). In 1954, a prominent gastroenterologist, Palmer, publishes a study that appears to demonstrate that no bacteria is capable of colonizing the human stomach. Palmer looks at biopsies from over 1,000 patients and observes no colonizing bacteria. As a result, he concludes that all previous observations of bacteria were a result of

[3]Klebs found bacteria in the gastric glands in 1881 (Fukuda et al., 2002), Jaworski observed these bacteria in sediment washings in 1889 (Kidd and Modlin, 1998), Bizzozero observed spiral organisms in dogs in 1892 (Figura and Bianciardi, 2002), and Saloon found similar spirocetes in the stomachs of cats and mice in 1896 (Buckley and O'Morain, 1998).

[4]Although, bismuth (an antimicrobial) had been used to treat ulcers dating as far back as 1868, the first report of an antibiotic occurs in 1951 (Unge, 2002).

contamination (Palmer, 1954).

The result of this study was the widespread abandonment of the bacterial hypothesis, poetically described by Fukuda, et al.,

> His words ensured that the development of bacteriology in gastroenterology would be closed to the world as if frozen in ice... [They] established the dogma that bacteria could not live in the human stomach, and as a result, investigation of gastric bacteria attracted little attention for the next 20 years (Fukuda et al., 2002, 17-20).

Despite this study, a few scientists and clinicians continued work on the bacterial hypothesis. John Lykoudis, a Greek doctor, began treating patients with antibiotics in 1958. By all reports he was very successful. Despite this, he was unable to either publish his results or convince the Greek authorities to accept his treatment. Undeterred, he continued using antibiotics, an action for which he was eventually fined (Rigas and Papavassiliou, 2002). Although other reports of successful treatment with antibiotics or observation of bacteria in the stomach occasionally surfaced, the bacterial hypothesis was not seriously investigated until the late 1970s.

At the 1978 meeting of the American Gastroenterology Association in Las Vegas, it appeared that tide had begun to turn. At this meeting, it was widely reported that the current acid control techniques could not cure ulcers but merely control them (Peterson et al., 2002). When antacid treatment was ceased, the symptoms would invariably return. The very next year, Robin Warren first observed Helicobacters in a human stomach, although reports of this result will not appear in print until 1984 (Warren and Marshall, 1984).

Initial reactions to Warren and Marshall's discovery were negative, primarily motivated by the widespread acceptance of Palmer's conclusions. Marshall became so frustrated with his failed attempts to convince the scientific community of the relationship between *H. pylori* and PUD, that he drank a solution of *H. pylori*. Immediately after, he became ill and was able to cure himself with antibiotics (Marshall, 2002). Eventually after replication of Marshall and Warren's studies, the scientific public became convinced of Palmer's error. It is now widely believed that *H. pylori* causes PUD, and that the proper treatment for PUD involves antibiotics.

While we may never really know if Palmer engaged in intentional misconduct, the facts clearly suggest that he did not.[5] Palmer failed to use a silver stain when investigating his biopsies, instead relying on a Gram stain. Unfortunately, *H. pylori* are most evident with silver stains and are "Gram negative", meaning they are not easily seen by using the Gram stain. Although the silver staining technique existed in the 1950s, it would have been an odd choice for Palmer. That stain was primarily used for neurological tissue and other organisms that should not be present in the stomach. Warren did use the silver stain, although it is not clear what lead him to use that stain.

Although we cannot look into the souls of each and every scientist working on PUD, one can hardly criticize their behavior. They became aware of a convincing study, carefully done, that did not find bacteria in the stomach. Occasionally, less comprehensive reports surfaced, like those of Lykoudis, but since they contradicted what seemed to be much stronger evidence to the contrary, they were dismissed. Had the acid theory turned out to be true, the behavior of each individual scientist would have been laudable.

Despite the fact that everything was "done by the book," so to speak, one cannot resist the urge to think that maybe things could have been done differently. In hindsight, Palmer's study was too influential. Had it not been as widely read or been

[5] Marshall speculates that the long delay between the reports of his discovery and the widespread acceptance of the bacterial hypothesis were (partially) the result of the financial interests of pharmaceutical companies (Marshall, 2002). While this may be an example of pathological science, the dismissal of the bacterial hypothesis from 1954-1985 certainly is not.

as convincing to so many people, perhaps the bacterial theory would have won out sooner. Thinking just about PUD, these are likely idle speculations. On the other hand, we might think about the phenomenon of good faith errors more generally. We might then ask, what features of individual scientists and scientific communities might help to make these communities less susceptible to errors like Palmer's.

2 A First Model

In order to discuss this situation more generally, we will turn to abstract models of scientific behavior. Through the analysis of these models, perhaps we can gain insight into ways that scientific communities may be more and less reliable to errors of this sort.

Consider a stylized circumstance, superficially similar to the situation faced by PUD researchers. There are four medical researchers working on a particular disease. They are confronted with a new method of treatment which might be better or worse than the current, well-understood, method of treatment. Work on the new treatment will help to determine whether it is superior. Since the old treatment is well understood, work on it will not result in any new information about its probability of success, scientists' efforts will only refine delivery methods or reduce harmful side-effects. Suppose our scientists, labeled A, B, C, and D, assign the following probabilities to the superiority of the new treatment: 0.33, 0.49, 0.51, and 0.66. They then each pursue the treatment method which they think best. Two scientists, C and D, will pursue the new treatment option and two, A and B, the old. Suppose, further that the new treatment is in fact better than the old but, as is perfectly possible, C and D's experiments both conclude slightly against it. Specifically suppose all agree on these probabilities:

$$
\begin{aligned}
P(\text{The result of } C\text{'s experiment} \mid \text{New method is better}) &= 0.4\\
P(\text{The result of } D\text{'s experiment} \mid \text{New method is better}) &= 0.4\\
P(\text{The result of } C\text{'s experiment} \mid \text{New method is worse}) &= 0.6\\
P(\text{The result of } D\text{'s experiment} \mid \text{New method is worse}) &= 0.6
\end{aligned}
$$

After meeting and reporting their results to each other A, B, C, and D now asses the probability of the new theory being better as 0.1796, 0.2992, 0.3163, and 0.4632 respectively. As a result, none of them will pursue the new treatment; we have lost a more beneficial treatment forever. This outcome is far from extraordinary; given that the new methodology is better and the experimental outcomes are independent (conditioned on the new methodology being superior), the probability of getting this result is 0.16.

This circumstance arises for two reasons. First, scientists in our example must pursue evidence, they are not passive observers. Second, they already have a good understanding of the old treatment and further study of it will not help them to conclude anything about the new treatment.[6]

Even given this structure, the availability of the evidence contributes to the abandonment of the superior theory. Had D not been aware of C's result, however, she would still have believed in the superiority of the new treatment.[7] As a result, had she been unaware of C's results, she might have performed a second round of experiments, and had a chance to convince others of the benefits of the new treatment. In this toy

[6] Had the scientists been passive observers, their beliefs would not have influenced the type of information they received. In that case, information about either treatment might still arrive despite the fact that the theory has been abandoned. Additionally had experiments on the old theory been informative about the effectiveness of the theory, the fact that everyone pursues the old theory does not preclude them from learning about the new theory.

[7] If D had only been aware of her own negative results, but not the results of C, her posterior belief in the superiority of the new treatment would have been 0.5621.

example, it seems that the wide availability of experimental results was detriment to the group's learning. Of course, no general lesson can be draw from this example. It is not offered as a general model for all scientific practice, but is instead provided as a generalization of a learning situation that some scientists unquestionably face.

Two economists, Bala and Goyal (1998) present a very general model that can be applied to circumstances like the one faced by the medical researchers. Stated formally, in this model, there are two states of the world ϕ_1 and ϕ_2 and two actions A_1 and A_2. Action A_1 has the same expected return in both states while A_2's is lower in ϕ_1 and higher in ϕ_2. Agents are aware of the expected payoff in both states, but are unaware of which state obtains. Agents have beliefs about the state of the world and in each period take the action which has the highest expected utility given their beliefs. They receive a payoff from their actions which is independently drawn for each player from a common distribution with the appropriate mean. Each agent observes the outcome of his actions and the outcome of *some* others, and then updates his beliefs based on simple Bayesian reasoning about the state of the world.[8]

This model has multiple interpretations, but one of them is analogous to the circumstance discussed above. The agents are scientists and their action is choosing which method to pursue. ϕ_1 and ϕ_2 respectively represent the state where the current method and the new method is better. Bala and Goyal endeavor to discover under what conditions correct convergence can be guaranteed. They consider two different restrictions, restrictions on priors and restrictions on information.

The restriction on priors will be discussed in more detail in Section 3, in the context of a different model. The second suggestion, limiting information, will be our primary focus here. This restriction is achieved by limiting which other agents

[8] "Simple" here means that the agent only updates her belief using the evidence from the other's experiment. She does not conditionalize on the fact that her counterpart performed a particular experiment (from which she might infer the results of others).

an individual can "see," and thus restricting the information on which an agent can update. They do this by placing an agent on a graph and allowing her only to see those agents with which she is directly connected.

Bala and Goyal consider agents arranged on a line where each agent can only see those agents to the immediate left and right of them. If there are an infinite number of agents, convergence in this model is guaranteed so long as every agents' prior is in the interior of the probability space.[9] Bala and Goyal also consider adding a special group of individuals to this model, a "royal family." The members of the royal family are connected to every individual in the model. If we now consider this new collection of agents, the probability of converging to the wrong result is not zero! This is a remarkable result, because it contradicts a basic intuition about science: that access to more data is always better.[10] In this case, it is not.

The reason for this result is interesting. In the single line case the probability that everyone receives misleading results becomes vanishingly small as the population grows to infinity. However, in the population with the royal family, this probability no longer vanishes. Negative results obtained by the royal family infect the entire network and mislead every individual. Once the entire population performs act A_1, they can no longer distinguish between the good and bad states because this action has the same expected payoff in both ϕ_1 and ϕ_2. As a result a population composed entirely of A_1 players will never escape.

Interpreting this model in the case of peptic ulcer disease, one might see Palmer as an instantiation of the royal family. His results, were read by everyone. As a result

[9] The interior of the probability space is any value other than 0 or 1.

[10] Ellison and Fudenberg (1995) present a different model which comes to the same conclusions. In their model, the interaction structure is not fixed, individual take a different random sample of fixed size in each time period. Because, the individual in their model have much shorter memories, it seems less appropriate for modeling scientific behavior (an application which they do not consider). As similar conclusion can be found for even individual learning in the work of Herron et al. (1997). This work presents a rather different learning situation and will not be discussed here.

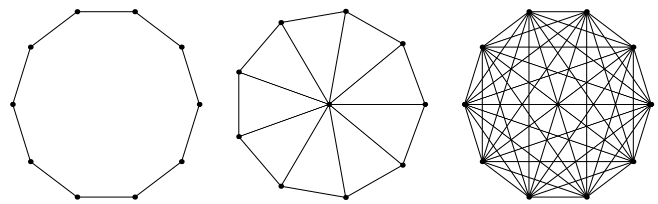

Figure 1: A 10 person cycle, wheel, and complete graph

they all abandoned the the bacterial theory, and turned to controlling acid. Had they not all been aware of his results, some of them may have continued investigating the bacterial theory. In that case, they might have found (more quickly) evidence in favor of the bacterial theory, and perhaps convinced others to ignore Palmer's results.

One might worry about Bala and Goyal's results since they depend so critically on the infinite size of the population. For finite populations, there exists a positive probability that *any* population will not converge. One might wonder, in these cases how much influence the "royal family" would have on the population. Furthermore, it is unclear what moral we ought to draw from these results – many things are different in the two different models. In addition to increased connectivity, there is also unequal distribution of connections. If we are interested in evaluating the performance of actual institutions it is unclear which features we should seek out. Through computer simulations, we will endeavor to discover the influence that network structure has on reliable learning in finite populations and also to develop more detailed results regarding the relationship between network structure and success.

2.1 Finite Populations

2.1.1 The "Royal Family" Effect

To begin, we will look at three graphs known as the cycle, the wheel, and the complete graph (pictured in Figure 1) and compare their convergence properties. The cycle is a finite analogy to Bala and Goyal's line. Here agents are arranged on a circle and only connected with those on either side of them. The wheel is a cycle but one of the agents is connected to everyone else, Bala and Goyal's royal family. The last network is one where everyone is connected to everyone.

We will, unbeknownst to our agents, set the world in ϕ_2, where the new methodology is better. We will then assign our agents random beliefs uniformly drawn from the interior of the probability space and allow them to pursue the action they think best. They will then receive some return (a "payoff") that is randomly drawn from a distribution for that action. The agents will then update their beliefs about the state of the world based on their results and the results of those to which they are connected. A population of agents is considered finished learning if one of two conditions are met. First, a population has finished learning if every agent takes action A_1, in this case no new information can arrive which will convince our agents to change strategies. (Remember that the payoff for action A_1 is the same in both states, so it is uninformative.) Alternatively the network has finished learning if every agent comes to believe that they are in ϕ_2 with probability greater than 0.9999. Although it is possible that some unfortunate sequence of results could drag these agents away, it is unlikely enough to be ignored.

The results of a computer simulation are presented in Figures 2 and 3. In Figure 2, the x-axis represents total number of agents and y-axis represents the proportion of 10,000 runs that reached the correct beliefs.[11] The absolute probabilities should not be taken too seriously as they can be manipulated by altering the expected payoffs for A_1 and A_2. On the other hand, the relative fact is very interesting. First, we have demonstrated that Bala and Goyal's results hold in at least some finite populations.

[11]Although it is possible for a population to continue unfinished indefinitely, no population failed to converge.

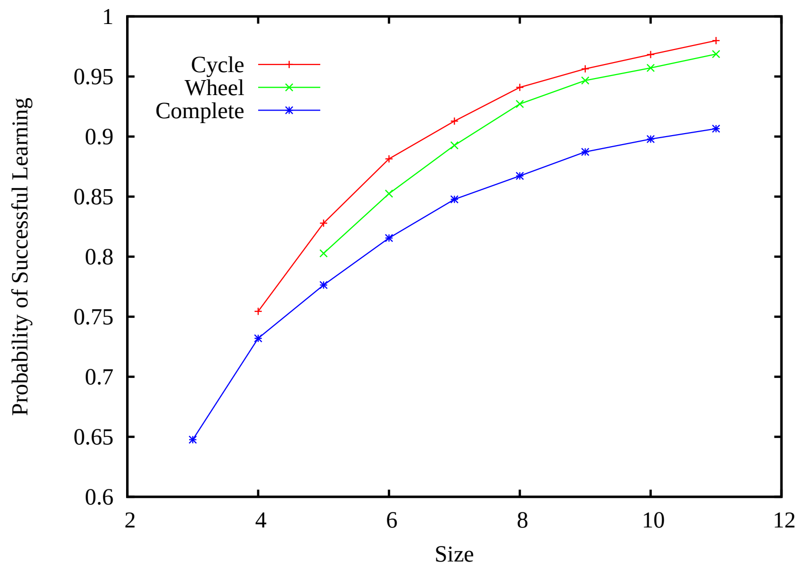

Figure 2: Learning results of computer simulations for the cycle, wheel, and complete graphs

In all the sizes studied the cycle does better than the wheel. Second, we have shown that both of these do better than the complete graph where each agent is informed of everyone else's results.

This demonstrates a rather counterintuitive result, that communities made up of less informed scientists might well be more reliable indicators of the truth than communities which are more connected. This also suggests that it is not the unequal connectivity of the "royal family" that is the culprit in these results. The harm done by the individual at the center cannot be simply overcome by removing their centrality.

There is a benefit to complete networks, however; they are much faster. Figure 3 shows the average number of generations it takes to reach the extreme beliefs that constituted successful learning among those networks that did reach those beliefs. Here we see that the average number of experimental iterations to success is much lower for the complete network than for the cycle, and the wheel lies in between. This

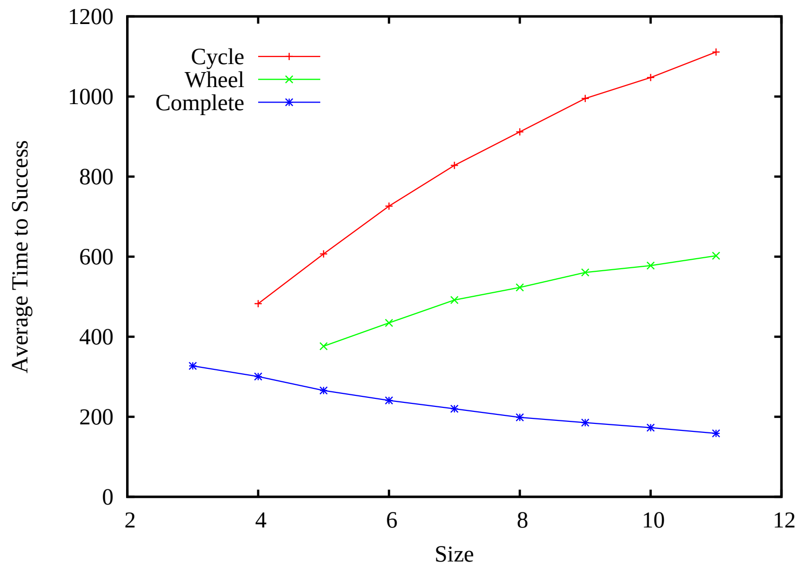

Figure 3: Speed results of computer simulation for the cycle, wheel, and complete graphs

suggests that, once networks get large enough, a sacrifice of some small amount of accuracy for the gain of substantial speed might be possible.[12]

So far, we have only looked at the properties of three networks, the trend seems to be that increased connectivity corresponds to faster but less reliable convergence. This is generalizing from three, relatively extreme networks, however. It would be good to engage in a more systematic survey.

2.1.2 Connectivity and Success

For relatively small sizes (less than seven) we can exhaustively search the properties of all networks. The suggestion in the previous section, that decreased connectivity results in slower, but more reliable learning, can be tested more extensively. In the

[12]The results for both reliability and speed are robust for these three networks across modifications of both the number of strategies (and thus states) and the difference in payoff between the good and uninformative actions. Although these different modifications do effect the ultimate speed and reliability of the models, for any setting of the parameters the relationship between the three networks remains the same.

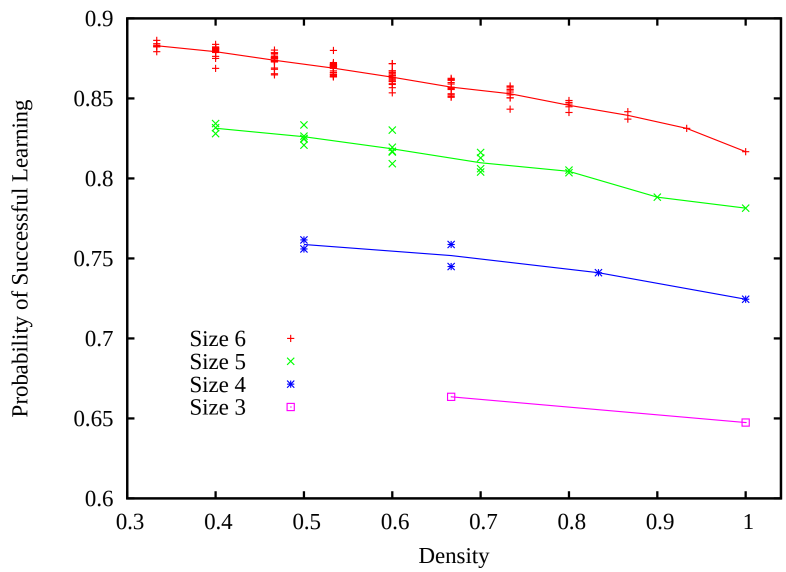

Figure 4: Density versus probability of successful learning

previous section, connectivity was left as an intuitive criterion. In fact, there are several graph statistics that correspond to our notion of connectivity. Here, we will use *density* which represents the percentage of possible connections that actually obtain in a graph.

Taking all networks (up to isomorphism) between size three and six we can compare these statistics to network's learning properties. These results are presented in Figure 4.[13] A regression among the largest group (networks of size six) reveals that density is a stronger predictor of successful learning than any other common graph statistic. Only one other graph statistic significantly improves the prediction beyond density alone, that is the *clustering coefficient*.[14] This statistic measures the degree to which one's neighbors (those to whom an individual is connected) are connected to each other. In both cases, the lower the statistic (i.e. the less dense and less

[13]This is a result of running 10,000 trials for every graph, up to isomorphism, of size 6 or lower. The initial beliefs were independently drawn from a uniform distribution over [0,1].

[14]In fact, there are two clustering coefficients in the literature. One from Newman et al. (2002) is a slightly better predictor than one first presented in Watts (1999) although the difference is very small.

clustered a graph is) the higher the successful learning. In addition, the in-network degree variance is not correlated with success, suggesting that it is not the centrality of the wheel, but its high connectivity that results in its decrease in reliability.[15]

Examining the differences among the different finite cases is instructive. It appears that sparsely connected networks have a much higher "inertia." This inertia takes two forms. First, an unconnected network experiences less widespread change in strategy on a given round than a highly connected network. The average number of people who change their strategies after the A_2 players receive less than expectation is four times higher in a highly connected network than a less connected network. Second, unconnected networks are less likely to occupy precarious positions than connected ones. Conditioning on the network having only one A_2 player, a highly connected network is almost three times as likely to have no individuals playing A_2 on the next round. Since there is only one new piece of evidence in both cases, the difference between the two networks is the result of individuals having less extreme beliefs (i.e., closer to 0.5) in the connected network. Since all networks have the same expected initial beliefs, this must be the result of the information received by the agent.[16]

Both of these results suggest that unconnected networks are more robust to the occasional string of bad results than the connected network because those strings are contained in a small region rather than spread to everyone in the network. This allows the small networks to maintain some diversity in behaviors that can result in the better action ultimately winning out if more accurate information is forthcoming. This also explains why we observed the stark difference in speeds for the cycle and

[15] The in-network degree variance measures the variability of the number of neighbors for each agent in a network. This number will be high when certain agents are connected to more people than the average, thus representing unequal connectivity. Since this statistic is uncorrelated with success, we can conclude that the centrality of the royal family does not influence its reliability.

[16] The statistics reported here are comparing 100 runs of a complete six person networks with the most reliable six person network pictured in Figure 5.

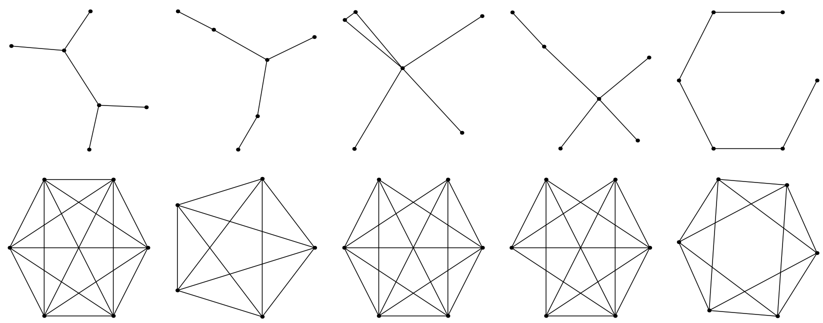

Figure 5: The five most accurate (top) and five fastest (bottom) networks

complete networks in the previous section. When bad information is contained so too is good information. In fact, we find that this trade off is largely robust across networks.

An inspection of the five most reliable and five fastest networks suggests that the features of a network that make it fast and those that make it accurate are very different (see Figure 5). Four of the five most reliable graphs are minimally connected – i.e., one cannot remove any edge without essentially making two completely separate graphs. Conversely, the five fastest graphs are highly connected, two of them are complete graphs, and the remaining ones are one, two, and three edges removed from complete graphs. Figure 6 compares the average time to success and probability of success for networks of size six. Here we find that there is a relationship between the accuracy of a network and its speed. In fact, this graph shows that sometimes a small increase in probability can result in a substantial increase in time to success.

This confirms the tradeoff suggested before, in order to gain the reliability that limiting information provides, one must sacrifice other benefits, in this case, speed. In fact, the tradeoff is even stronger than suggested here. These results are only for cases where we specify that the new method is better. When the uninformative action is better convergence is guaranteed but the connectedness of the graph determines its

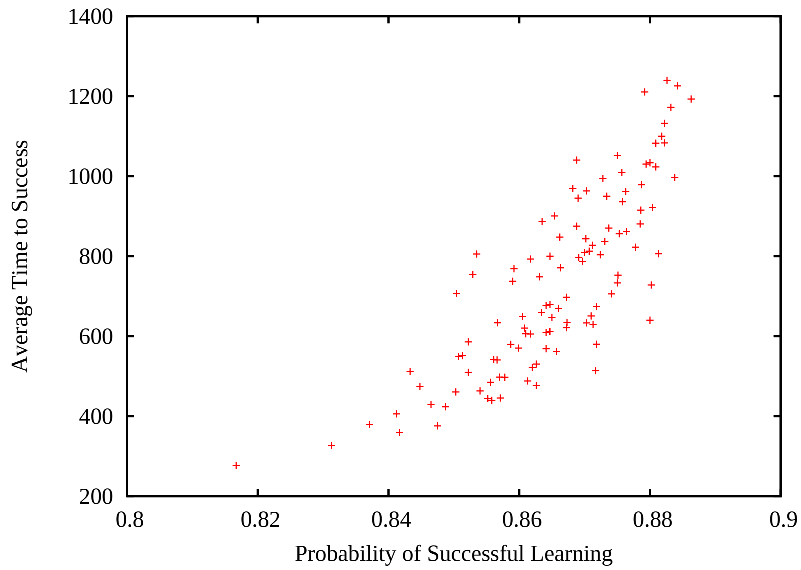

Figure 6: Speed versus accuracy for networks of size six

speed.

In the previous section, relationship between speed and size was a strange one. For complete networks, as the network grew the average time to success of these groups decreased. On the other hand, for wheels and cycles as the network grew the average time to success increased. This diversity is verified by the more complete analysis. There appears to be no correlation between size and speed when all networks are considered.

Ultimately, there is no right answer to the question of whether speed or reliability is more important – it will depend on the circumstance. Although a small decrease in reliability can mean a relatively large increase in speed, in some cases, such sacrifices may not be worth making. If it is critical that we get the right result no matter how long it takes we would prefer groups where information is limited (without making the network disconnected). On the other hand, if speed is important and correct results are not as critical perhaps a more connected network is desired. It is not the intention of this study to provide unequivocal answers to these questions, but rather to demonstrate that such trade-offs do exist and that one can achieve increased reliability by limiting information.

2.2 The Right Model

There are three assumptions that underlie this model which might cause some concern. They are:

1. The learning in our model is governed by the observation of payoffs.

2. There is a uninformative action whose expected payoff is well known by all actors.

3. The informative action can take on one of very few expected payoffs and the possibilities are known by all actors.

The first assumption is of little concern. Here we use payoffs to symbolize experimental outcomes. Payoffs that are closer to the mean are more likely, which corresponds to experimental outcomes that are more likely on a given theory. The payoffs are arranged so that an individual who maximizes her expected payoff pursues the theory that she thinks is most likely to be true. This fact allows this model to be applied to learning situations where individuals are interested in finding the most effective theory (however effectiveness is defined) and also to situations where individuals are interested in finding the true theory. In either case the individuals behave identically.[17]

The second and third assumptions are less innocuous. In the next section we will study a model where these assumptions are relaxed. However, while this study is important, it should not be presumed that the Bala-Goyal model is in some way

[17] This is not to say that true theories always have higher payoffs. Instead, this model is so general as to apply to either circumstance.

deficient. In fact, this model very closely mimics Larry Laudan's (1996) model of theory selection.

Laudan suggests that theory choice is a problem of maximizing expected return. We ought to choose the theory that provides the largest expected problem solving ability. Since we have often pursued a particular project for an extended time before being confronted with a serious contender, we will have a very good estimate of its expected utility. However, we will be less sure about the new contender, but we could not learn without giving it a try.

Even beyond Laudan, there may be particular learning circumstances that conform to these three assumptions. Bala and Goyal compare their model to crop adoption in Africa. There, a new seed is introduced and farmers must decide whether to switch from their current crop (whose yield is well known) to another crop (whose yield is not). Experimental techniques and apparatus may well follow a similar pattern.

There are, however, many learning situations where these assumptions do not hold. The case of PUD presents one example. When Palmer conducted his study, neither the bacterial theory nor the acid theory was widely accepted. The effectiveness of antibiotics had not been widely studied. In this case, the scientists had no knowledge of the possible benefits of the two competing research programs beyond some commonsensical upper and lower bounds. Nonetheless scientist were put in a position where they were forced to pursue one or the other. We would be interested if, in learning situations of this type, the same qualitative facts about reliable networks remains true.

3 Beta Learning Model

In the previous section we looked at a model where agents are learning between two alternative states of the world. In that model the payoff to pursuing different theories was uncorrelated because one theory had the same payoff in both states, while the other varied between the two states. There we assumed that the agents knew the available options for the expected payoffs and were trying to learn which one was the true one.

In some circumstances this may be a plausible model of scientific practice. When our background theories give us a good reason to suppose that one theory is either slightly better or slightly worse than another, well understood, theory, this model applies. However, there are also many cases when we have no idea about the payoffs for different theories. As a simple example suppose we have two coins and we are trying to learn which coin has a higher probability of coming up heads.[18] Suppose further that there is no a priori reason to believe that the probabilities are correlated.

There are many different ways that we might model Bayesian agents learning the probability of a coin. One of the most natural is to use beta distributions.

3.1 Beta distribution

In the case of the coin flipping example there are an infinite number of possible hypotheses of the form $p_i = x$ where $x \in [0, 1]$ and i represents the individual coin. As a result we cannot model this using the simple Bayesian model applied in Section 2. There is more than one way that one might want to model Bayesian learning of continuous values. For instance, one might divide the continuum into finitely many

[18] In more scientific terms, suppose that heads represents a successful application of a scientific theory. Here we want to choose the coin that comes up heads most often, i.e., the scientific theory that is most applicable.

intervals and have the agents learn on each of the intervals. A better approach (both computationally and theoretically) is to use beta distributions.

Beta distributions are a class of probability density functions defined on the interval $[0, 1]$. Since the range is continuous, we cannot assign a non-zero probability to "$p_i = x$" for any specific x without assigning an infinite number of x's probability 0. Instead, we opt to use probability density functions (pdf). Beta distributions represent one class of pdfs.

Definition 1 (Beta Distribution) *A pdf on* $[0, 1]$*,* f*, is a beta distribution iff for some* α *and* β

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}$$

Where $B(\alpha, \beta) = \int_0^1 u^{\alpha-1}(1-u)^{\beta-1}du$.

The beta distribution has some very nice properties. In addition to having only two free parameters it has the property that if one has a beta distribution as a prior, and one takes a sample of any size and updates, one will have a beta distribution as a posterior. Suppose one performs n flips of the coin and receives s heads. If an individual has a beta distribution with parameters α and β as a prior, then their posterior will also be a beta distribution with the posterior parameters $\alpha + s$ and $\beta + n - s$ (cf. DeGroot, 1970). This makes the beta distribution a *conjugate prior.*

The expectation for a beta distribution is given by $\frac{\alpha}{\alpha+\beta}$. This enables a rather brief convergence result. Suppose an agent starts out with priors α and β. She then performs a series of trials which result in s successes in n trials. Her posterior is has parameters $\alpha + s$ and $\beta + n - s$. As a result her posterior mean is given by:

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

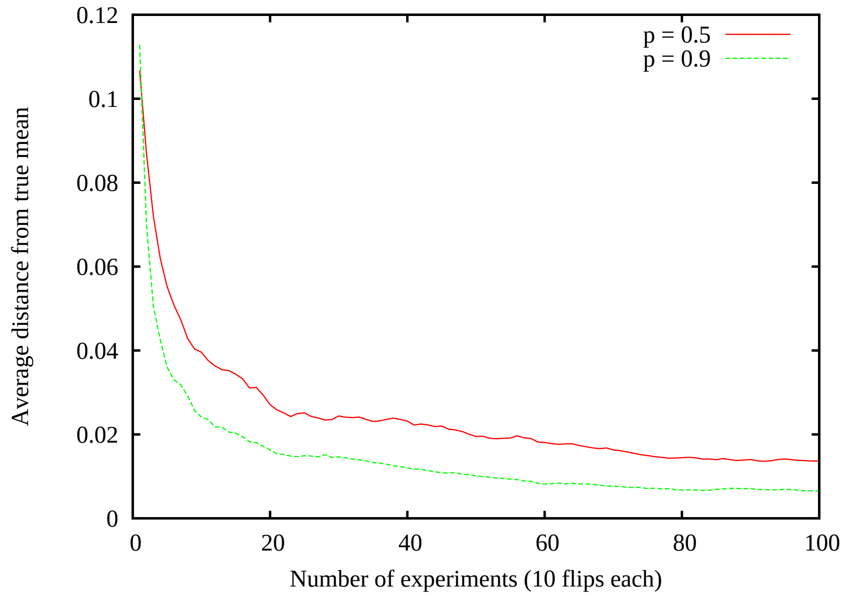

Figure 7: Learning using beta distributions

As s and n grow, $\frac{s}{n}$ approaches the true probability of successes and since they will eventually grow well beyond α and β, the mean of the agent's beliefs will approach the true probability (cf. Howson and Urbach, 1996).

Unfortunately, proof of limiting convergence does not always give us the information we desire. Ellison (2000) analyzes a model where the limiting properties for a particular game are well defined, but the model can spend *arbitrarily long* in another state. Here it is not clear that the limiting properties correspond to anything we wish to study. If a system can spend arbitrary long in a state to which it does not converge, then its convergence properties neither provide satisfactory predictions nor normative recommendations.

Luckily Bayesian learning with beta distributions is relatively efficient. Figure 7 illustrates simulation results for a single agent learning the probability of a coin. She starts with random initial parameters (randomly drawn from a uniform distribution on $[0, 4]$) and flips the coin ten times. After this, the agent updates her priors. The x-axis represents the number of such experiments that have been performed. The lines represent the average distance from the true mean for 100 such individuals. Here we see that beta distribution learning can be very fast. Interestingly it appears that beta learning is faster when the probability of success is more extreme. With this version of Bayesian learning in hand we can turn to the model.

3.2 The model

As in Section 2 individuals will confront a choice between two competing actions. They will take an action (e.g. pursue developments in a particular theory), which will return a number of successes and failures based on an underlying probability. One action will have a higher probability of success than the other. The agents will update their beliefs about the expectation of each action on the basis of their results and on the results of their neighbors.

Individuals will be assigned a random initial α_i and β_i for each action i from the interval $[0, 4]$. They will then take the action which has the highest expectation. As in the previous model, failed learning is possible. Consider a single learner case, where the individual has the following priors:

$$\begin{aligned}
\alpha_1 &= 1 \\
\beta_1 &= 3 \\
\alpha_2 &= 3 \\
\beta_2 &= 1
\end{aligned}$$

This yields an expectation of 0.25 for action 1 and 0.75 for action 2. Suppose that action 2 does have an expectation of 0.75, but action 1 has a higher expectation of 0.8. Since the individual thinks that action 2 is the superior action, he will take it on the first round. So long as the results he receives do not take him too far from the expectation of action 2, he will continue to believe that action 2 is superior and will never learn that his priors regarding action 1 are skewed.

This type of failed learning is not as pernicious as in the last model. There, once an agent had developed incorrect beliefs, no possible outcome could dissuade her – she had converged. This was the result of one action being uninformative, it had the same payoff in all states. Here failure is not as strong. In the example above, there are strings of possible outcomes that will drag our agents expectation for action 2 below 0.25. If one of these strings of outcomes occurs, the agent will then try action 1, and, perhaps, this will result in the posterior expectation of action 1 increasing. On the other hand, this does not guarantee correct learning in finite time, because there are a number of results which will not bring the expectation for action 2 so low.

Because of this possibility, the definition of success and failure in learning must change. Since for any setting of the parameters there is a possible string of outcomes that will cause a different action to become preferred, mere uniformity of action does not guarantee success. Instead, we will allow our individuals to go for a set number of trials and at the end, record whether there is diversity in action, and if not whether they are all taking the best or worst action.

Figure 8 shows the results for the cycle, wheel and complete graphs.[19] Again we see the same relationship that held in the previous model. Unlike the results in Figure 2, the differences here are much more pronounced. A few differences are worth noting. First, in the previous model as the number of agents grew all three networks improved at approximately the same rate. Here, the complete network is improving, but only slightly. Second, the distance between the three networks is much larger,

[19]Each trial represents 1,000 flips of one of the coins, which have a .5 and .499 probability of coming up heads respectively. This is identical to the parameters used for the model presented in Section 2.

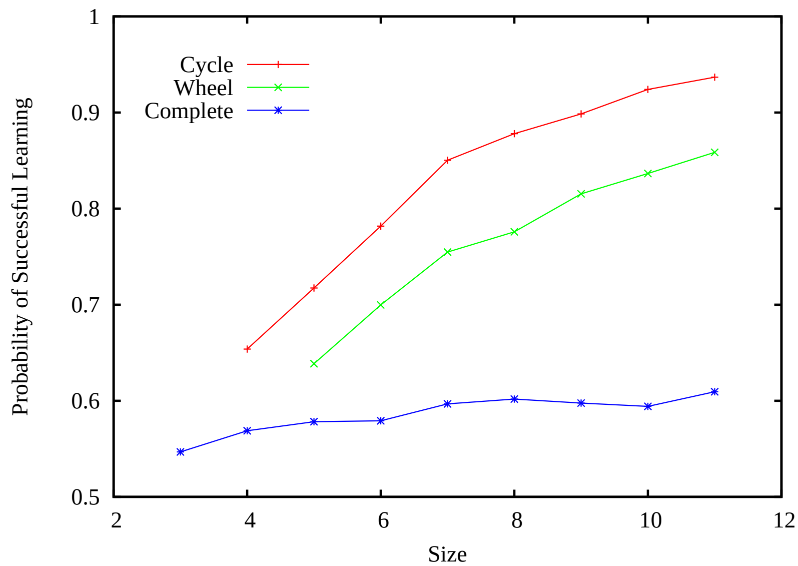

Figure 8: Results for the cycle, wheel, and complete graph

even at small sizes.

As in the previous model we can again look at all networks up to size six. The same qualitative relationships hold, the more connected a graph is the less reliable it becomes. Interestingly, a regression reveals correlations to several different statistics which measure connectivity. Again, however, all of these bear the same relationship to reliability.

3.3 Different priors

In beta distributions the size of the initial α's and β's determines the strength of an individuals prior belief. Figure 9 illustrates three different beta distributions with differing parameters. Although all three distributions have the same mean, the ones with higher initial parameters have lower variances. Also, these distributions will be more resistant to initial changes. For instance if all three receive 8 successes in 12 trials, the expectations after updating will be 0.6, 0.56, and 0.52. The more extreme the initial beliefs an individual has, the more resistant to change she is.

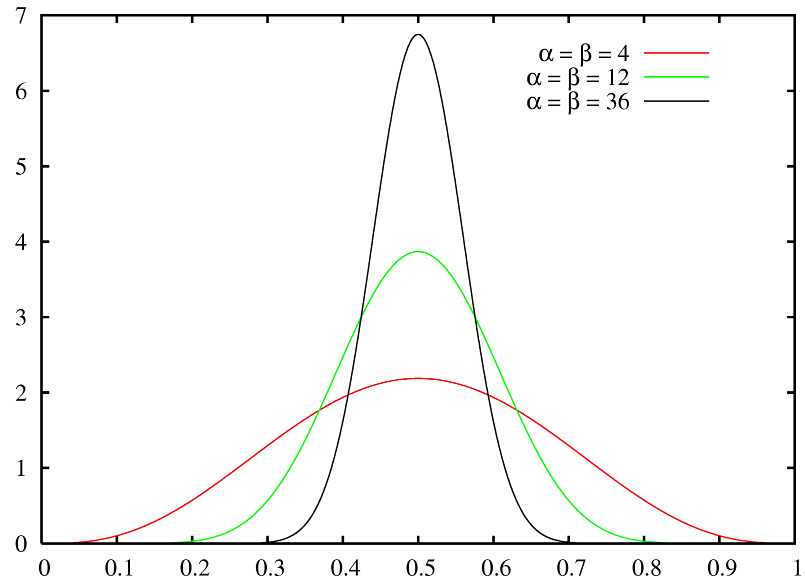

Figure 9: Three beta distributions

Since failed learning is a result of misleading initial results infecting the entire population, one might suggest that increasing an individual's resistance to change might help to alleviate the dangers of failed learning. In a more informal setting Karl Popper suggested this possibility. "A limited amount of dogmatism is necessary for progress. Without a serious struggle for survival in which the old theories are tenaciously defended, none of the competing theories can show their mettle" (1975, 87).[20] If this where the case, it might turn out that the difference between less connected and more connected networks would vanish.[21]

In order to investigate this possibility we will observe three canonical networks (a seven person cycle, wheel, and complete graph) and vary the range of initial beliefs. The results in the previous section were for α and β values between zero and four. Figure 10 shows the results for these three graphs as the maximum possible α and β

[20]A similar sentiment is echoed by David Hull (1988, 32).

[21]Bala and Goyal consider just such a suggestion the context of our first model. They find that for sufficiently extreme priors, convergence can be guaranteed for any network.

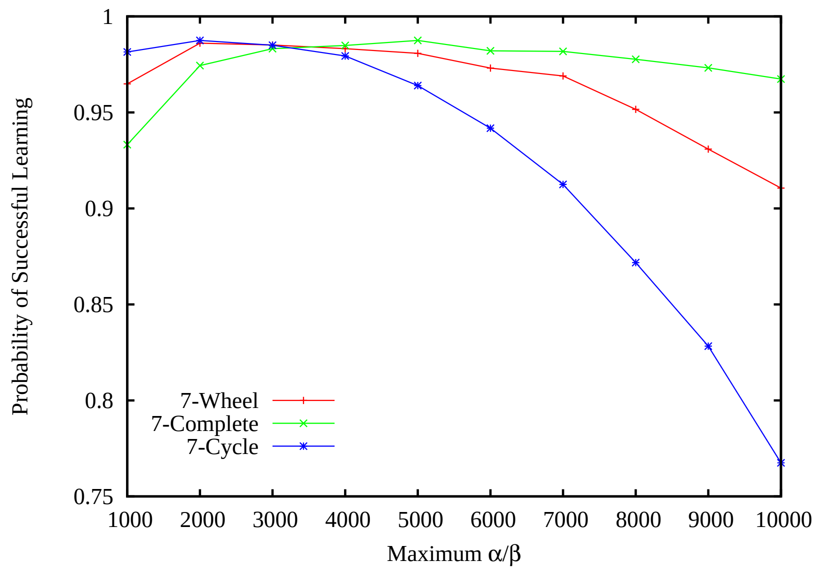

Figure 10: Success versus maximum α and β

value is increased. Individuals can still have very low initial parameters, but as the maximum increases this becomes less likely.

The results are quite striking. For α's and β's drawn from $[0, 1000]$, the results are similar to smaller initial parameters. However, as these parameters grow, the order of the networks reverses itself. At very extreme initial parameters, the complete network is by far the best of the three networks. This is not simply the result of one network reducing its reliability, but rather one network gains while the other loses.

The cause of this reversal is interesting. In the model from Section 2, complete networks were worse because they learned too fast. The wealth of information available to the agents sometimes caused them to discard a superior action too quickly. In the more limited networks this information was not available and so they did not jump to conclusions. This resulted in a longer time to success but ensured that optimal actions were not discarded too quickly.

Similarly, when our agents have very extreme priors, even rather large amounts of information will not cause them to discard actions. As a result, the benefit to

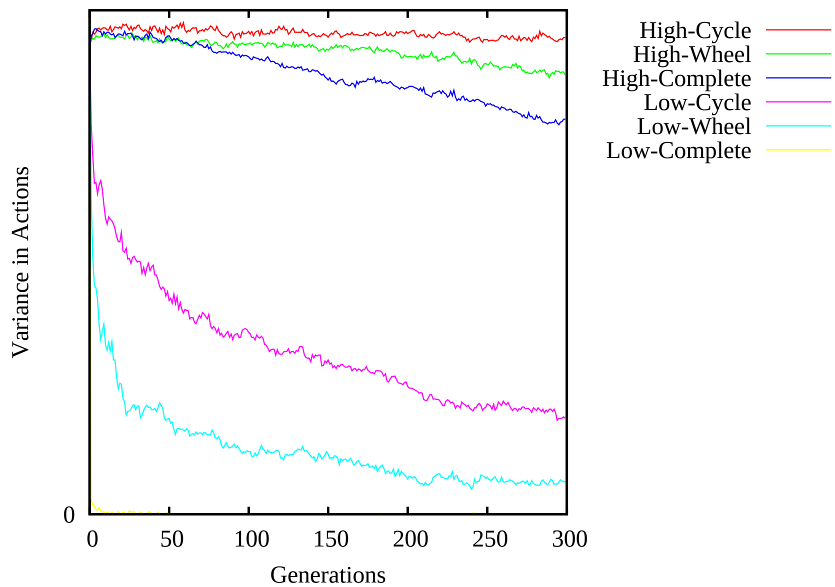

Figure 11: In network variance over time

disconnected networks vanishes – no matter how connected the network, agents will not discard theories too quickly. This explains why the complete network is not so bad, but does not explain why the less connected network becomes worse. In the less connected network there is simply not enough information to overcome the extremely biased priors within sufficient time. Since the simulations are stopped after a certain number of trials, the extreme priors have biased the agents sufficiently that they cannot overcome this bias in time. In fact, the overall frequency of failed learning did not increase in these networks. Instead the drop in reliability of the less connected networks was a result of a substantial increase in networks that failed to finish learning, either successful or not.

All of this is illustrated in Figure 11. Here the top three networks are a seven person cycle, wheel, and complete network with very extreme priors.[22] The bottom three are the same but with priors drawn from a much smaller distribution. The

[22]In this case alpha and beta are drawn from a uniform distribution on [0,7000].

y-axis represents the mean variance of the actions taken on that round. The higher the variance, the higher, on average, the diversity in actions.

The steep drop off of the three networks with limited priors is to be expected. The fact that the cycle and wheel preserve their diversity for much longer than the complete network illustrates the benefits of low connectivity. (In fact, the complete network drops of so quickly that is it almost invisible on the graph.) All of the networks with extreme priors maintain their diversity much longer, however, the complete network begins to drop of as more information accumulates.

An interesting feature discovered by varying the extremity in the priors can be seen in Figure 10. Around 3000 the three networks almost entirely coincide. Here, we are slightly worse off than at either of the extremes (cycle with unbiased priors or complete with extreme priors). On the other hand, this represents a sort of low risk position, since the network structure is largely irrelevant to the reliability of the model.

Returning again to the case of PUD, if individual scientists had been more steadfast in their commitment to the bacterial hypothesis they might not have been so convinced by Palmer's study. Perhaps they would have done studies of their own, attempting to find evidence for a theory they still believed to be true. Because of their steadfast commitments we would want all of the information to be widely distributed so that eventually, we could convince everyone to abandon the inferior theory once its inferiority could be clearly established.

4 Conclusion

This last result illustrates an important point. At the heart of these models is one single virtue, transient diversity. This diversity should be around long enough so that individuals do not discard theories too quickly, but also not stay around so long as to hinder the convergence to one action. One way of achieving this diversity is to limit the amount (and content) of information provided to individuals. Another way is to make individuals' priors biased. However, each of these are not independent virtues. Both together make the diversity too stable, and result in a worse situation than either individually.

For PUD, we have suggested that things might have been better had Palmer's result had not communicated so widely or had people been so extreme in their beliefs that they remained unconvinced by his study. However, it would have been equally bad had both occurred simultaneously. In the actual history, 30 years were wasted by pursuing a sub-optimal treatment. Had the scientists been both uninformed and dogmatic, we might still be debating the bacterial and hypoacidity hypothesis today.

The benefit of diversity in beliefs (and actions) in scientific communities is not new with this study. Philip Kitcher (1990; 1993; 2002) and Michael Strevens (2003a; 2003b) have looked at a different set of models, where diversity has some benefit. Like the models presented here, in Kitcher and Streven's models diversity is needed in order to ensure that potentially beneficial scientific theories are not discarded. Although the details of these models differ from the one considered here, the underlying motivation is the same. Unlike Kitcher and Strevens, this study provides a series of solutions to the diversity problem which they do not consider. In addition, this study articulates a negative consequence of diversity, a problem which does not occur in Kitcher and Streven's models.

These solutions each turn on individuals being arranged in ways that make each individual look epistemically sub-optimal. The scientists do not observe all of the available information or have overly extreme priors. Looking at these scientists from the perspective of individualistic epistemology, we might criticize the scientists' behavior. However, when viewed as a community, their behavior becomes optimal. This confirms a conjecture of David Hull, who says:

> Quite obviously science is a social process, but it is also "social" in a more significant sense. The objectivity that matters so much in science is not primarily a characteristic of individual scientists, but of scientific communities.... The mechanism that has evolved in science that is responsible for its unbelievable success may not be all that "rational," but it is effective, and it has the same effect that advocates of science as a totally rational enterprise prefer (1988, 3-4)

Limiting information or dogmatic priors have a good effect when the overall behavior of the community is in focus. This suggests that when analyzing particular behaviors of scientists (or any epistemic agents) we ought to think not just about the effect their behavior has on their individual reliability, but on the reliability of the community as a whole.

References

Bala, V. and S. Goyal (1998). Learning from neighbours. *Review of Economic Studies 65*, 565–621.

Buckley, M. J. and C. A. O'Morain (1998). Helicobacter biology - discovery. *British Medical Bulletin 54*(1), 7–16.

DeGroot, M. H. (1970). *Optimal Statistical Decisions*. McGraw-Hill.

Ellison, G. (2000). Basin of attraction, long-run sotachastic stability and the speed of step-by-step evolution. *Review of Economic Studies 67*, 17–45.

Ellison, G. and D. Fudenberg (1995, February). Word-of-mouth communication and social learning. *The Quarterly Journal of Economics 110*(1), 93–125.

Figura, N. and L. Bianciardi (2002). Helicobacters were discovered in Italy in 1892: An episode in the scientific life of an eclectic pathologist, Giulio Bizzozero. In B. Marshall (Ed.), *Helicobacter pioneers: Firsthand accounts from the scientists who disocovered Helicobacters*, pp. 1–13. Victoria, Australia: Blackwell Science Asia.

Fukuda, Y., T. Shimoyama, T. Shimoyama, and B. J. Marshall (2002). Kasai, Kobayashi and Koch's postulates in the history of Helicobacter pylori. In B. Marshall (Ed.), *Helicobacter Prioneers*, pp. 15–24. Blackwell publishers.

Goldman, A. (1999). *Knowledge in a Social World*. Oxford: Clarendon Press.

Herron, T., T. Seidenfeld, and L. Wasserman (1997, September). Divisive conditioning: Further results on dilation. *Philosophy of Science 64*, 411–444.

Howson, C. and P. Urbach (1996). *Scientific Reasoning: The Bayesian Approach* (2nd Edition ed.). Chicago: Open Court.

Hull, D. (1988). *Science as a Process*. Chicago: University of Chicago Press.

Kidd, M. and I. M. Modlin (1998). A century of Helicobacter pylori. *Digestion 59*, 1–15.

Kitcher, P. (1990, January). The division of cognitive labor. *The Journal of Philosophy 87*(1), 5–22.

Kitcher, P. (1993). *The Advancement of Science*. New York: Oxford University Press.

Kitcher, P. (2002). Social psychology and the theory of science. In S. Stich and M. Siegal (Eds.), *The Cognitive Basis of Science*. Cambridge: Cambridge University Press.

Laudan, L. (1996). *Beyond Positivism and Relativism: Theory, Method, and Evidence*. Boulder: Westview Press.

Marshall, B. (2002). The discovery that Helicobacter pylori, a sprial bacterium, caused peptic ulcer disease. In B. Marshall (Ed.), *Helicobacter pioneers: First hand accounts from the scientists who discovered Helicobacters*, pp. 165–202. Victoria, Australia: Blackwell Science Asia.

Newman, M., D. Watts, and H. Strogatz (2002). Random graph models of social networks. *Proceedings of the National Academy of Sciences 99*(1), 2566–2572.

Palmer, E. (1954). Investigations of the gastric mucosa spirochetes of the human. *Gastroenterology 27*, 218–220.

Peterson, W. L., W. Harford, and B. J. Marshall (2002). The Dallas experience with acute Helicobacter pylori infection. In B. Marshall (Ed.), *Helicobacter pioneers: firsthand accounts from the scientists who discovered Helicobacters*, pp. 143–150. Victoria, Australia: Blackwell Science Asia.

Popper, K. (1975). The rationality of scientific revolutions. In R. Harre (Ed.), *Problems of Scientific Revolution: Progress and Obstales to Progress*. Oxford: Clarendon Press.

Rigas, B. and E. D. Papavassiliou (2002). John Lykoudis: The general practitioner in Greece who in 1958 discovered the etiology of, and a treatment for, peptic ulcer disease. In B. Marshall (Ed.), *Helicobacter Pioneers*, pp. 75–87. Blackwell publishers.

Strevens, M. (2003a). Further properties of the priority rule. *Manuscript*.

Strevens, M. (2003b). The role of the priority rule in science. *Journal of Philosophy 100*(2), 55–79.

Unge, P. (2002). Helicobacter pylori treatment in the past and in the 21st century. In B. Marshall (Ed.), *Helicobacter pioneers: Firsthand accounts from the scientists who discovered Helicobacters*, pp. 203–213. Victoria, Australia: Blackwell Science Asia.

Warren, J. R. and B. J. Marshall (1984, June 16). Unidentified curved bacilli on gastric epithelium in active chronic gastritis. *Lancet 1*(8390), 1311–1315.

Watts, D. (1999). *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton: Princeton University Press.