Miguel, Amblard, Barceló & Madella (eds.) Advances in Computational Social Science and Social Simulation
Barcelona: Autònoma University of Barcelona, 2014, DDD repository <http://ddd.uab.cat/record/125597>

On Fragmentation and Scientific Progress

Stefano Balietti, Michael Maes and Dirk Helbing
ETH Zurich, Switzerland
<sbalietti@ethz.ch> <mmaes@ethz.ch> <dhelbing@ethz.ch>

1 Introduction

Why are the social sciences so fragmented as compared to the natural sciences, for example physics or biology? What consequences has fragmentation for the overall progress of a field?

The focus of this work is the question why disciplinary fragmentation and scientific progress are correlated. Potential explanations are abound. First, there might be a direct causal relationship between fragmentation and scientific progress. For instance, it could be that the natural sciences are less fragmented into opposing schools because they have developed a scientific consensus very early in their history (Cole, 2001, Chap. 6). Likewise, a high degree of fragmentation might slow down scientific progress, because it hampers the diffusion of ideas and insights (Cole, 2001, Chap. 6). Second, fragmentation and progress might not influence each other, but there might be third variables that affect both outcomes. These factors might be rooted in the nature of the disciplines. For instance, many concepts of social scientific theories are difficult to define, leading to disagreement about their appropriate interpretation and measurement (Cole, 2001, Chap. 1). Unclear definitions of concepts also slow down the process of scientific discovery. Furthermore, institutional factors might affect the degree of fragmentation and progress in the different disciplines (Cole, 2001, Chap. 14). For instance, scarcity in public funding does not only slow down progress but might also further increase competition between scientist and hamper the willingness to interact with competing research teams, which in turn fosters the formation of distinct clusters.

2 Model Specification

We assume a population of $N = 100$ scientists exploring a continuous epistemic space under the effect of social influence of related opinions. The position $\vec{x}_i(t)$ that scientist i occupies in the epistemic space at time t represents his or her current *view*. At time $t + 1$, scientist i can either update or not his position. Such a movement in space then represents the *theoretical approach* that the agent follows in his or her investigation. We specify the exploration of the epistemic landscape by the scientist in search for the correct theory as follows:

$$\frac{d\vec{x}_i}{dt} = \underbrace{\vec{v}_i(t)}_{Exploration\ velocity} + \underbrace{\vec{C}_h(t)}_{Convergence} + \underbrace{\vec{\xi}_i(t)}_{Noise} . \quad (1)$$

The *exploration velocity* term is specified as in Vicsek et al. (1995) model for collective motion of self-propelled particles:

$$\vec{v}_i(t+1) = \alpha \vec{v}_i(t) + (1-\alpha)\langle \vec{v}_j(t) \rangle_R , \quad (2)$$

where $0 < \alpha \leq 1$ and $\langle . \rangle_R$ means an averaging over all scientists within the radius R. This specification of the exploration speed reflects the desire to produce scientific progress along with one's peers, but also some persistence in the research path taken.

However, unlike the original Vicsek's model, the velocity is not a constant, but varies over time. In fact, each scientist is initially assigned a velocity $\vec{v}_i(t_0)$ drawn from a uniform random distribution. The averaging over neighbors eventually produces an alignment of the velocities vectors of both components of Eq. 2, e.g. the direction and the strength of the vector. Given that agents that meets at a certain position can follow completely antithetical approaches, the size of the velocity vector is usually reduced by the averaging procedure. This represents frictional effects in social interactions.

The *convergence term* $\vec{C}_h(t)$ reflects the attraction to the ground truth. As mentioned also by Hegselmann and Krause (2006), it is not to be interpreted literally, but rather as the feedback that individuals receive from the results of their experiments and continued investigation. The force is suppose to give a weak hint regarding the research direction to follow, but never the exact answer, because otherwise every individual could jump directly to the true value, and this is simply not how scientific research typically develops. Formally $\vec{C}_h(t)$ is modeled as:

$$\vec{C}_h(t) = \frac{\vec{x}_h - \vec{x}_i(t)}{\tau_h}\,, \tag{3}$$

where $\vec{x}_h - \vec{x}_i(t)$ is the distance between the agent's position and the location of the ground truth, and τ_h is a parameter representing the strength of the attractive force. It implies that initially, when still far away from the ground truth, scientists receive a clear signal of the research direction they should take to approach the truth. However, the closer they get to the true value, the harder it gets to actually make progress.

The noise term $\vec{\xi}_i(t)$ plays a crucial role in our model in a twofold manner. On the one side, we assume (i) Gaussian position noise on the approach hold by a scientist at time t, and on the other (ii) we assume Gaussian angular noise on the research direction followed by each individual scientist at time $t+1$. The position noise, whose standard deviation is denoted by ϵ, acts as white noise, while the angular noise, whose standard deviation is denoted by σ, introduces stocasticity and path dependence in the trajectory followed by the scientist. The former can be interpreted as an measurement error, which can be reduced but never completely eliminated. The latter can instead be interpreted as the inevitable difference between planned and actual research path, or as well as the degree to which a scientist is voluntarily ready deviate from the approach of his or her peers, if he or she is making research in a group.

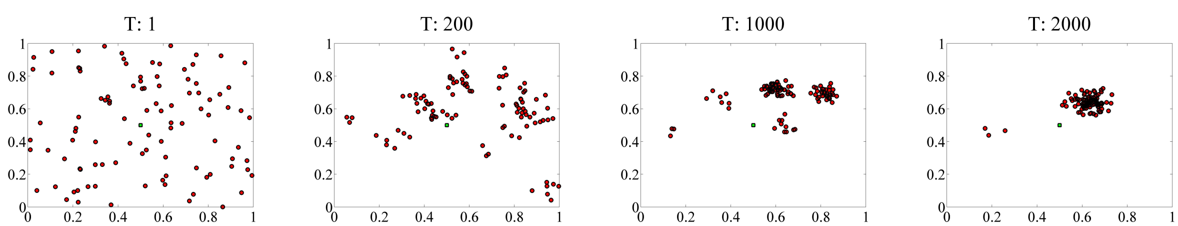

Figure 1: Ideal typical run showing different patterns of consensus and level of scientific progress at four time steps: 1, 200, 1000, 2000. The figure shows the emergence of different schools of thought, that gradually merge together in proximity of the truth.

3 Preliminary Results

We devised the following computational experiment. At the beginning of each simulation we have randomly preassigned agents to $n = (1, \ldots, 30)$ clusters. Each group of agents was then placed on a radius of 0.4 units away from truth, and equidistant from neighboring groups. Afterward, we have then measured the time necessary to achieve a stable *consensus* on truth, defined as 75% of the agents within a radius of 0.05 units from truth. The results are shown in Fig. 2 A. Unequivocally, we observe two different consensus regimes, determined by the value of the interaction radius R. In the case of small R, there is a significant positive relationship between the number of clusters and the time necessary for reaching consensus. In the case of large R, the number of cluster has a very limited effect. Surprisingly, however, the coefficient even takes the opposite sign: the more clusters the faster the consensus is built up. The results from the linear regression of time to consensus over initial number of clusters are shown in Table 1.

We then investigated, the effect of the other social influence variable of our model, i.e. the strength of social influence α. In case of a large radius R, as expected, manipulating α played no major role. However, in case of a small radius R, the results are striking. When social influence is very strong ($\alpha = 0.01$), the R^2 value of the regression is 0.51. On the contrary, if social influence is very weak ($\alpha = 0.99$), no simulation reached consensus within 20.000 iterations!

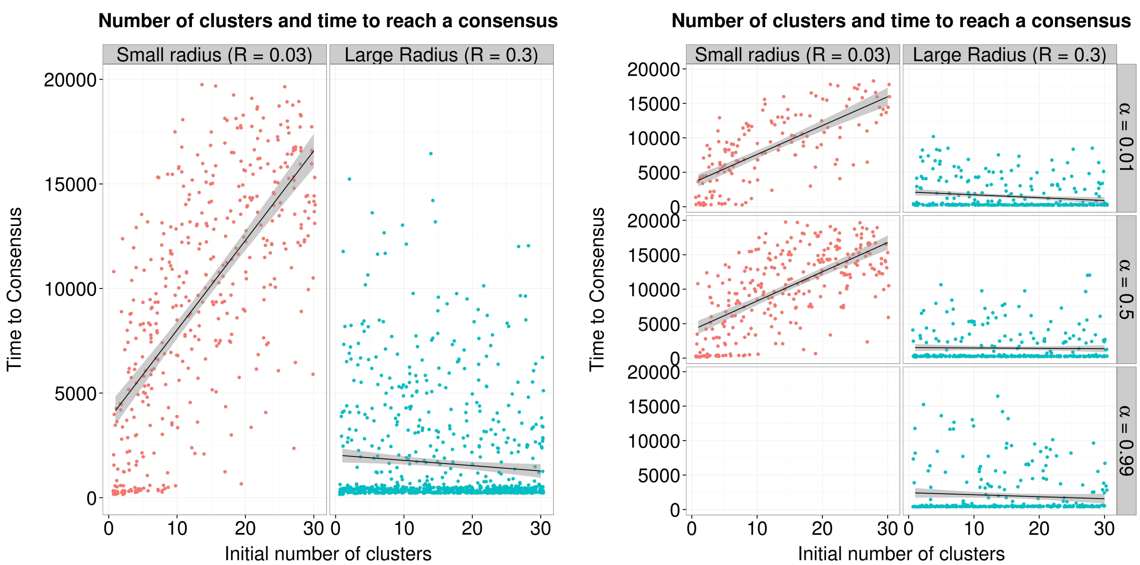

Figure 2: At the beginning of the simulation agents were randomly preassigned to $n = (1, \ldots, 30)$ clusters. Each group of agent was then placed on a radius of 0.4 units away from truth, and equidistant from their neighboring groups. The graph shows the time necessary to 75% of the agents to end up in a radius of 0.05 units from truth. Clusters are a hurdle to consensus if and only if the interaction radius is small. On top of this, if also social influence is also weak, e.g. $\alpha = 0.99$, reaching the truth can be extremely unlikely.

Clustering per se is not a factor that slows down scientific progress. It is rather the nature of social interactions that completely determines how consensus is built up and how long it takes to reach the ground truth. A small interaction radius, and highly individualistic agents represent the worst environmental conditions to form a consensus, but also to find the truth.

Even if clustering does not directly retard scientific progress, it is in fact a proxy for identifying the real cause: the small radius of influence used by the agents in the field. Large radius, even in connection

	Model 1	Model 2	Model 3	Model 4
(Intercept)	3735.33^{***}	2037.75^{***}	3406.91^{***}	2142.94^{***}
	(364.17)	(176.41)	(482.20)	(250.77)
Clusters	428.33^{***}	-25.31^{*}	418.27^{***}	-41.39^{**}
	(22.37)	(9.94)	(32.95)	(14.13)
R^2	0.46	0.01	0.51	0.03
Num. obs.	435	900	155	300

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 1: Model **1**: $alpha = [0.01; 0.5; 0.99]$, and $R = 0.03$. Model **2**: $alpha = [0.01; 0.5; 0.99]$, and $R = 0.3$. Model **3**: $alpha = 0.01$, and $R = 0.03$; Model **4**: $alpha = 0.01$, and $R = 0.3$;

with low social influence, do not generally lead to a clustered field. The confounding nature of clustering and social influence has is such that fragmentation has often been erroneously seen as the primary cause of the slow rate of progress of disciplines like sociology. Our simulation results clearly show that clustered fields can make quick progress if agents are open to social influence, and use a large enough radius to combine the results of investigations of others with their own.

References

Berger, R.L. (1981). "A Necessary and Sufficient Condition for Reaching a Consensus Using DeGroot's Method". In: *Journal of the American Statistical Association* 76.374, pp. 415–418.

Cole, S. (2001). *What's Wrong with Sociology?* Transaction Publishers.

De Langhe, R. (2012). "Peer Disagreement Under Multiple Epistemic Systems". In: *Synthese*, pp. 1–10.

DeGroot, M.H. (1974). "Reaching a Consensus". In: *Journal of the American Statistical Association* 69.345, pp. 118–121.

Feyerabend, P. (1994). "Art as a Product of Nature as a Work of Art". In: *World Futures* 40, pp. 87–100.

Fortunato, S. et al. (2005). "Vector Opinion Dynamics in a Bounded Confidence Consensus Model". In: *International Journal of Modern Physics C* 16.10, pp. 1535–1551.

Harary, F. (1959). "Studies in Social Power". In: ed. by D. Cartwright. Univer. Michigan. Chap. A Criterion for Unanimity in French's Theory of Social Power, pp. 168–182.

Hegselmann, R. and U. Krause (2002). "Opinion Dynamics and Bounded Confidence Models. Analysis and Simulation". In: *Journal of Artificial Societies and Social Simulation* 5, pp. 1–24.

— (2006). "Truth and Cognitive Division of Labour: First Steps Towards a Computer Aided Social Epistemology". In: *Journal of Artificial Societies and Social Simulation* 9.3, pp. 1–28.

Kitts, J.A. and P. T. Trowbridge (2007). "Shape Up Or Ship Out: Social Networks, Social Influence, and Organizational Demography". In: *Computational and Mathematical Organization Theory* 13, pp. 333–353.

Lehrer, K. (1975). "Social Consensus and Rational Agnoiology". In: *Synthese* 31.1, pp. 141–160.

Maes, M., A. Flache, and D. Helbing (2010). "Individualization as Driving Force of Clustering Phenomena in Humans". In: *PLoS Comput Biol* 6.10, e1000959+.

March, J.G. (1991). "Exploration and Exploitation in Organizational Learning". In: *Organization Science* 2.1, pp. 71–87.

McPherson, M., L.S. Lovin, and J.M. Cook (2001). "Birds of a Feather: Homophily in Social Networks". In: *Annual Review of Sociology* 27.1, pp. 415–444.

Mouzelis, N. (1995). *Sociological Theory: What Went Wrong?: Diagnosis and Remedies*. Routledge.

Muldoon, R. (2013). "Diversity and the Division of Cognitive Labor". In: *Philosophy Compass* 8.2, pp. 117–125.

R.P., Abelson (1964). "Contributions to Mathematical Psychology". In: ed. by N. Frederiksen and H. Gulliksen. New York: Rinehart Winston. Chap. Mathematical Models of the Distribution of Attitudes under Controversy, pp. 142–160.

Sandstrom, P.E. (1999). "Scholars as Subsistence Foragers". In: *Bulletin for the American Society of Information Science* 25.3, pp. 17–20.

Turner, S. and J.H. Turner (1990). *The Impossible Science: An Institutional Analysis of American Sociology*. Sage.

Vicsek, T. et al. (Aug. 1995). "Novel Type of Phase Transition in a System of Self-Driven Particles". In: *Physical Review Letters* 75.6, pp. 1226–1229.

Weinberg, S. (1998). "Physics and History". In: *Daedalus* 127.1, pp. 151–164.

Weisberg, M. and R. Muldoon (2009). "Epistemic Landscapes and the Division of Cognitive Labor". In: *Philosophy of Science* 76, pp. 225–252.

Whitley, R. (2000). *The Intellectual and Social Organization of the Sciences*. Oxford University Press.

Zollman, K. J.S. (2010). "The Epistemic Benefit of Transient Diversity". In: *Erkenntnis* 72.1, pp. 17–35.