Scientific collaboration: do two heads need to be more than twice better than one?

Thomas Boyer-Kassem and Cyrille Imbert
Archives H. Poincaré (CNRS and Université de Lorraine, UMR 7117, Nancy, France)

to appear in *Philosophy of Science*

Preprint version: June 2015

Abstract

Epistemic accounts of scientific collaboration usually assume that, one way or another, two heads really are more than twice better than one. We show that this hypothesis is unduly strong. We present a deliberately crude model with unfavorable hypotheses. We show that, even then, when the priority rule is applied, large differences in successfulness can emerge from small differences in efficiency, with sometimes increasing marginal returns. We emphasize that success is sensitive to the structure of competing communities. Our results suggest that purely epistemic explanations of the efficiency of collaborations are less plausible but have much more powerful socio-epistemic versions.

1 Introduction.

Scientific collaboration has been continually developing since the 19th century (Beaver and Rosen 1978, 1979a, 1979b). More and more articles are now produced collaboratively, by larger and larger teams. In 2000, only 1 article out of 5 was authored by a single scientist and the average collaboration size was 4 (Wuchty, Jones and Uzzi 2007)[1]. Clearly, collaboration results from various factors which vary across fields and situations. Such factors can be non-epistemic like citation artifacts or power relations, or epistemic, like a higher research quality. The underlying idea of many of these explanations of collaboration is that, as expressed by Thagard (2006, 194), "two heads really are more than twice better than one"—otherwise, why collaborate?

We shall show in this paper that, surprisingly, that much is not actually needed. We propose a deliberately crude model in which collaboration boils down to the sharing of intermediate results and investigate how such collaborative groups succeed in a competitive environment. The unexpected result we find is that, despite our conservative hypotheses, collaboration is by far more efficient, and small epistemic differences in efficiency can imply large differences in success. We shall further show that the successfulness of groups is sensitive to the collaboration structure of the community. Overall, our results imply both that purely epistemic accounts of collaborations are harder to defend and that incorporating social aspects make them more powerful.

The paper is organized in the following way. In section 2, we shall present existing accounts of scientific collaboration and introduce our modeling strategy. We shall then describe our model in section 3 and analyze it in section 4. Some general conclusions are drawn in section 5 before we discuss the robustness of our results in section 6 and return to the issue of the explanation of collaboration in section 7.

2 Existing strategies to account for collaboration.

We shall now present and characterize existing explanations of the actual development of scientific collaboration in the last two centuries and describe our modeling strategy.

First, collaboration can be explained by *non-epistemic* factors. For example, insofar as collaboration is a way to gain and sustain access to recognition in professional communities, it may "act as a social regulator" (Beaver and Rosen 1978, 69). Sometimes, it is emphasized that collaboration is pragmatically beneficial. For example, it may be a way of making one's ideas dominant (Latour 1987) or developing "powerful lobbying groups" (Wray 2002). It may be an effect of credit attribution, because publishing a paper with a co-author may bring a researcher almost the same amount of renown as the same single-authored paper might but requires less work[2].

Explanation of collaboration can also attribute an important role to *epistemic* factors. For example, the epistemic task may simply be too big to be achieved by one scientist during her lifetime (Hardwig 1985, 345–347). In particular, some accounts explicitly

[1] Collaboration is less significant in the social sciences and almost marginal in the arts and humanities (*ibid.*).

[2] Regarding the question of how to attribute authorship and credit, see Rennie, Flanagin and Yank 2000 and Wray 2006.

emphasize that collaboration is epistemically beneficial (hereafter EB-accounts). For example, collaboration may develop because collaborative researchers are more productive or reliable. These epistemic benefits of collaboration can be the results of epistemic processes (hereafter EB_E-accounts), non-epistemic processes (hereafter EB_{NE}-accounts) or a mixture of both (hereafter EB_{E+NE}-accounts)[3]. For example, explanations stating that collaboration produces more reliable research (if mistakes can be uncovered and corrected more easily) (Fallis 2006, 200) or generates valuable new ideas (Thagard 2006, 183) correspond to EB_E-accounts. In such cases, the underlying idea is that, in the end, "two heads really are (epistemically) more than twice better than one" and collaboration is worthwhile because at the task level, the efficiency of the group is super-additive in the sense that it exceeds the sum of the efficiencies of the individuals composing it.

EB_{NE}-accounts and mixed EB_{E+NE}-accounts show how social factors can be closely connected with epistemic success. Wray (2002) seems to provide a EB_{NE}-account when he claims that contemporary research requires abundant resources and accessing them is easier to collaborative groups, which makes them epistemically successful. Specialization looks like a mixed EB_{E+NE}-account since it is a partly social feature rooted in the epistemic possibility of dividing scientific labor and thus being more efficient (Anderson 2013, Beaver and Rosen 1978, 69-70, Muldoon forthcoming). EB_{NE}-accounts and EB_{E+NE}-accounts acknowledge the intertwined roles played by non-epistemic and epistemic features, as advocated by Kitcher (1990), Solomon (1994) and Goldman (1999). Since such roles are rooted in general epistemic features, they can have a wide scope but, as they incorporate social aspects, they can also account for specific differences between fields or historical periods. As we shall see, our model provides material to criticize pure EB_E-accounts and shows the potential power of some EB_{E+NE}-accounts.

While the cited qualitative analyses carry valuable insights, all explanatory factors are unlikely to have the same importance. Moreover, as is well-known, intuitive explanations can be misleading and lead to seemingly neutral factors being neglected. Formal approaches, based on the modeling of scientific communities, can be used to analyze various factors in more detail and depth. For example, the literature about the division of cognitive labor has fruitfully investigated the relation between the success of scientific communities and their structure (De Langhe 2014, Kitcher 1990, Muldoon and Weisberg 2011, Strevens 2003). One important assumption present in such works is that the success function primarily depends on the total number of researchers sharing a practice. However, as no difference is made between collaborators and competitors, such models are ill-suited to analyze collaboration as such. Accordingly, we shall propose a model which describes the competition between distinct collaborative groups within a community.

Models can be developed with different strategies. In order to make accurate predictions, rich realist models, which include most—if not all—relevant details, are appropriate. Another modeling strategy is to analyze highly idealized models, with little content (see Weisberg 2007, 642-644 for a survey). Such models provide different epistemic benefits. They are better-suited to investigate the impact of specific properties or mechanisms.

[3]Presumably, these EB-accounts eventually need to include non-epistemic factors to explain how epistemic success explains the increase in the number of collaborating scientists. A plausible hypothesis is that pragmatic rewards, like positions or research grants, accrue to epistemically successful scientists so that they have more funding and students.

Thereby, they develop our understanding of how phenomena emerge and help distinguish the individual contribution of each property to the total effect. Using such models as a starting point for an inquiry is methodologically sound since this may reveal that the target phenomenon can emerge from unexpected mechanisms. For instance, in Schelling's model (1971), residential segregation was found to emerge from among agents with a high threshold of tolerance for mixed neighborhoods. Finally, the corresponding explanations can have large scope, like explanations of universal properties in physics (Batterman 2002).

To analyze collaboration, we provide one such highly-idealized model and abstract away most features of scientific groups and their research environments, with the exception of two basic and general factors, namely the sharing of intermediate results within a group and the priority rule, "possibly the most distinctive feature of the social organization of science" (Strevens 2003, 55). This strategy is all the more appropriate since generating models in which collaboration is beneficial is not difficult. Indeed, there is currently a wealth rather than a shortage of explanations which build the beneficialness of collaboration in the features of groups by assuming that, in one way or another, the task efficiency of groups is super-additive. Accordingly, we deliberately did not make this assumption—in our model, the task efficiency of groups arises from the individual properties of their members. Therefore, if collaboration is beneficial in our model, it will *a fortiori* be so if additional (and more realistic) beneficial mechanisms are also part of the picture and this will be informational about how little is needed for collaboration to emerge as a beneficial practice.

3 The model and its rationale.

In 1958, a theoretical paper by Schawlow and Townes, which presented the principle of functioning of an optical maser (later known as the "laser"), triggered a race between the top optical laboratories of the time. Half a dozen teams were involved, with sometimes many co-workers. Surprisingly, in 1960, Maiman, a physicist working alone, managed to make the first laser work. The other teams then aimed to be first to build a particular kind of laser. In the USA, three teams were involved. First was the team led by Townes himself, who worked with potassium vapor and later considered ruby and cesium instead. Another team was composed of Javan, Bennett and Herriott and worked on a helium-neon laser. Finally, Sanders, an English physicist invited by Kompfner, considered a pure helium laser project. The three teams all belonged to the Bell laboratories and faced similar scientific problems in their common inquiry but worked independently. In December 1960, after months of work on the optical problems, gas properties and resonator of the system, Javan, Bennett and Herriott managed to obtain a coherent light coming out of their apparatus: they had built the first gas laser. This historical episode instantiates the main features of the model to follow. There is a well-defined and publicly known scientific goal. Scientists compete to reach it first, face similar problems and some choose to gather in teams, which do not change during the competition.

Here is our model[4]. n scientists try to solve a well-defined problem, composed of l successive steps (see fig. 1). Only the last step is publishable. Time is discrete and

[4]Boyer (2014) uses a similar model to investigate different issues (see footnote 7).

researchers each have an objective probability p per unit time to pass a step. If p is close to 1 (respectively 0), the steps of the problem are easy to solve (respectively hard). On average, it takes l/p units of time for a single agent to complete the task, and p can be seen as an advancement pace.

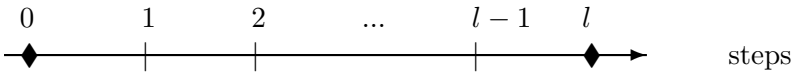


Figure 1: The scientific problem is sequential and composed of l steps.

The community of researchers abides by the priority rule—only the researcher who completes the final step first is rewarded. The reward (typically scientific credit) depends on the length and difficulty of the task. To make situations comparable, we set it to l/p, that is, proportional to the average time a single scientist needs to complete it. With this renormalization, scientists have no incentive to work on difficult or long problems, which does not particularly favor collaboration, since such problems are more likely to be solved by teams.

Researchers can either collaborate or not. Collaborating means that, if a researcher passes a step, she shares her achievement with her collaborators, who also pass this step. We suppose that the probabilities of any two researchers passing a step are independent[5]. This is coherent with the fact that researchers work independently and communicate only *once* a step has been passed. This in part favors collaboration since, if collaborators *always* succeeded or failed at the same time, a group would not succeed more than one single researcher. Overall, the probability that at least one out of k researchers passes a step is $p_g(k,p) = 1 - (1-p)^k$. This is the step efficiency of a collaborative group of k researchers (hereafter k-group). It is merely an aggregated effect with a sub-additive growth rate in k (since two researchers of group may pass a step simultaneously).

We assume that when a collaborative group gets the final reward, it is shared equally among collaborators. This hypothesis gives no extra-benefit to collaboration and might even undervalue how academic credit is distributed. For instance, for a job application, is a paper with three co-authors only credited as a third of a single-authored paper?

In brief, our model describes situations in which well-defined problems are tackled (so that knowledge of the research goal can be publicly known), results can be clearly identified by the community (so that the application of the priority rule makes sense) and the scientific task can be divided in some sequential well-identified steps. These conditions are often met in sciences like physics, biology, computer science, etc., fields in which collaboration is primarily found. Sharing intermediate results and the independence of research outcomes at the step level are the only factors favoring collaboration in our model. They do not make collaboration outstandingly beneficial, are general features of collaboration, and correspond to a possible (though simplistic) collaborative practice. Therefore, they correctly serve the purpose of choosing a conservative but sensible value for the step efficiency of groups. However, we do not claim that our model is minimal, or represents a worst-case scenario. As dysfunctional collaborations can be imagined, any negative hypothesis implying disadvantages for collaborators could be added to make

[5]This hypothesis is commonly made in similar modeling contexts (e.g. Boyer 2014, De Langhe 2014, Zollman 2007 and 2010).

collaboration less favorable (see also section 5.2). Nevertheless, it would not be coherent to include such disadvantages, if they are not tied to the collaborative mechanism described in the model. We shall come back to this issue in section 6. Overall, as far as the benefits of collaboration are concerned, our model is a conservative and unfavorable one. So, if collaboration turns out to be beneficial in *this* model, it will be a strong result.

4 Results: how beneficial is collaboration?

We shall now analyze how beneficial collaboration is in this model. Solving research problems is one thing but solving them quickly is another. Having a total of five publications is outstanding for a graduate student but not for a senior professor. Accordingly, the quantity which needs to be considered is reward *per unit time*. We define individual successfulness as the amount of reward an individual researcher gets (alone or by collaborating) divided by the time spent (this is the quantity that a selfish scientist wants to maximize) and total successfulness as the reward distributed to the n researchers divided by the time spent to get it. Therefore, total successfulness is the sum of the individual successfulnesses and maximizing this quantity is valuable for the community.

If scientists collaborate, they usually solve the problem more quickly than those working alone but the reward needs to be shared, so there is no simple answer to the question of whether collaboration is beneficial. We have studied the model using computer simulations[6], with large statistics (millions of runs) such that, on the forthcoming graphs, the statistical error bars would be smaller than what can actually be seen.

4.1 Case of $n = 2$ researchers. The case of two researchers is instrumental to understand the behavior of the model. To investigate it, we vary p from 0.1 to 1 and l from 2 to 1000 (e.g., if the temporal interval represents one week, with $p = 0.5$ and $l = 10$, a single researcher needs on average 20 weeks to complete the research project). The results are shown in fig. 2, which illustrates the influences of the ease (respectively length) of the task on the benefits that researchers get when they work alone or collaborate. Both graphs make it clear that individual successfulness is higher if researchers collaborate than if they do not. Therefore, it is in their interest to collaborate, and the longer or the more difficult the task is, the more beneficial collaboration is.

A qualitative argument may help understand this result. By symmetry, there is 1 chance out of 2 that a single researcher finishes first and gets the l/p reward. So, on average, researchers working alone get $l/(2p)$ but if they collaborate, their 2-group always gets the l/p reward (there are no other competitors) and, as the reward is split, they still get $l/(2p)$ each. However, successfulness corresponds to reward *per unit time*. When they collaborate, researchers advance quicker because when one collaborator passes a step, the whole group progresses. Therefore, on the whole, they get a higher reward per unit time. In this $n = 2$ case, the (only) effect of collaboration is to decrease the time needed to solve the problem. This *speedup effect* is responsible for the gap between the red and blue curves on each graph of fig. 2. Since total successfulness is the sum of individual successfulnesses, it is higher here if researchers collaborate, and this is also due to the speedup effect. The

[6] An analytic solution of the model may be reachable, especially for asymptotes, but, for the philosophical questions and range of values studied here, simulations are pragmatically as convenient.

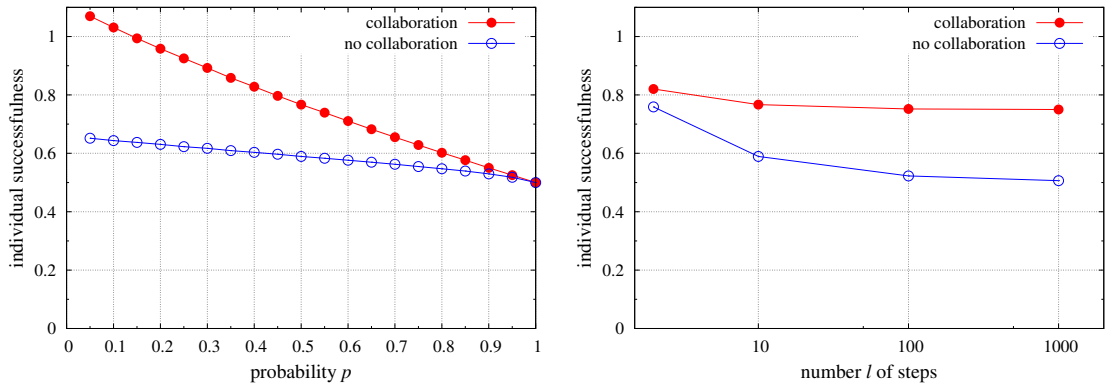


Figure 2: Influence of collaboration, for $n = 2$. Individual successfulness a) versus easiness of the task, with $l = 10$ (left) or b) versus length of the task, with $p = 0.5$ (right).

moral is that, here, it is both the interest of society and the two researchers to collaborate, whatever the difficulty or the length of the problem.

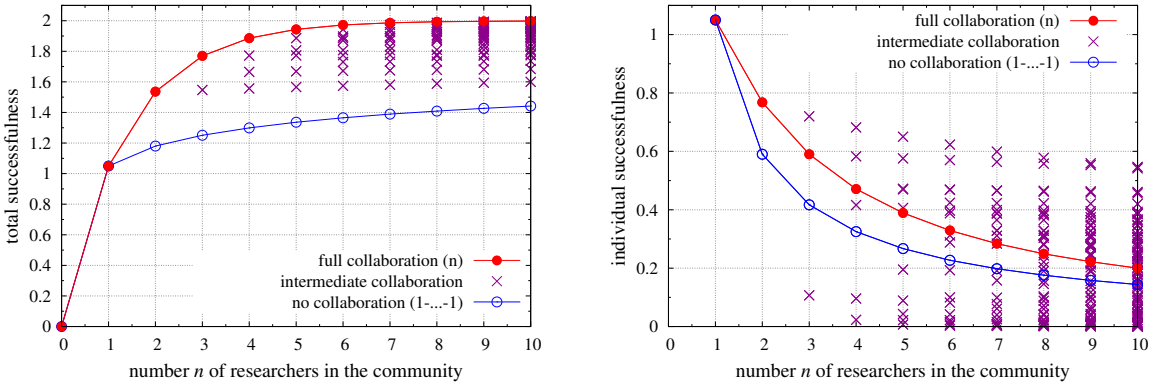
These results, beyond their apparent simplicity, represent a real departure from existing accounts of collaboration. While scientific collaboration is usually explained by assuming super-additive mechanisms that clearly favor groups, no such assumption is made here since collaborators simply share intermediate results and divide gains. Yet, these conservative hypotheses can be sufficient to make collaboration significantly beneficial for individuals (for long tasks, with $p = 0.5$ and $l = 100$, collaborators increase their successfulness by about 50%). Further, collaboration here emerges from an environment of fierce competition, since the priority rule is applied and, perhaps unexpectedly, is part of what makes collaboration beneficial.

One worry may be that the model is oversimplified since, so far, researchers should always collaborate. As we shall now see, for larger n, the model exhibits a more complex behavior.

4.2 Case of n up to 10 researchers. For $n = 3$, the possible configurations are a 3-group, a 2-group and a 1-group, or three single researchers—we denote these configurations (3), (2-1) and (1-1-1). The 138 configurations up to $n = 10$ were investigated by means of simulations to obtain the individual successfulness of researchers within the various k-groups and the total successfulness of the community. As the length or the difficulty of the problem do not change the results in qualitative terms, we set these at $p = 0.5$ and $l = 10$ steps.

Let us begin with the total successfulness of the researchers. Fig. 3a displays it as a function of n for all possible community configurations. Red points correspond to fully collaborative configurations (one n-group), blue points to non-collaborative configurations (n loners) and purple crosses to in-between configurations (e.g. for $n = 3$, the blue point stands for (1-1-1), the purple one for (2-1), and the red one for (3)). In fig. 3a, red points are the highest—because of the *speedup* effect, total successfulness is larger when researchers fully collaborate (this is actually true for any l and p). More generally, for

any collaboration configuration and any l and p, total successfulness increases when two groups merge. The reason is that the more researchers collaborate, the less time is lost by researchers in the community, by being stuck at steps passed by others. These results and their justification are similar to the ones given in Boyer (2014, section 4.1)[7]. Therefore, within this model, it is better for society that all researchers fully collaborate. Further, as can be seen with the full (respectively no) collaboration cases, the addition of new researchers always increases—though increasingly marginally—the total successfulness of a community. Indeed, at any time t, the chance that the task has been completed by at least one group is increased by the addition of a newcomer, either because she will be a lucky loner or because she will make her group more efficient.



(a) Total successfulness. Points stand for particular configurations of a community of size n.

(b) Individual successfulness. Points stand for particular k-groups in particular configurations.

Figure 3: Total and individual successfulnesses for all configurations ($p = 0.5$, $l = 10$).

Let us now consider the *individual* successfulness of each researcher. Fig. 3b displays it for all k-groups within all community configurations up to $n = 10$. (Since all members of a k-group have the same individual successfulness, we say in short that a point stands for a k-group in a given configuration.) For instance, for $n = 3$, the red point corresponds to the 3-group in (3), the green point corresponds to lone researchers in (1-1-1) and the higher (respectively lower) purple point to the 2-group (respectively loner) in (2-1). For each n, red points are higher than green, which means that, for individuals, full collaboration is better than no collaboration at all. But some purple points are still higher: for individuals, there exist intermediate configurations which are better than full collaborations (not for all the n members but at least for the researchers of some k-groups). For example, for $n = 3$, the most successful group is the 2-group in (2-1), which does slightly better than the 3-group in (3); lone researchers in (1-1-1) are doing fairly well, and the lone researcher in (2-1) poorly (see fig. 4)[8].

Overall, for $n \geqslant 3$, our model shows that some researchers can be better off not fully

[7]Our present model and Boyer's both consider a sequence of l steps, which agents pass with probability p. In Boyer (2014), any step can be published and the question for agents is whether they should publish now or later. When a result is published, the corresponding step is passed by all members of the community, whereas, in our model, intermediate results are shared with collaborators only. In both models, the more researchers share results (by collaborating or publishing), the less time is lost by other researchers, the

Group configuration	Individual successfulness	Total successfulness
(3)	0.590	1.770
(2-1)	0.720; 0.107	1.547
(1-1-1)	0.417	1.251

Figure 4: Individual and total successfulness with $n = 3$.

collaborating, whereas society would be better off if they did. This indicates a discrepancy between individual and collective interests (Kitcher 1990). This discrepancy might be seen as justifying the idea that society should encourage collaboration in scientific communities. Nevertheless, because other factors (like the need for diversity in methods) may pull in a different direction, treating this question in more depth would require going well beyond the scope of the present paper, so we shall not discuss it any further here.

To better understand why full collaboration is not necessarily the best option, we shall now analyze where the gain in successfulness for collaborative groups comes from. Consider for instance the shift from (1-1-1) to (2-1). The gain in *total* successfulness is 0.296 and it is due to the saving of time in the 2-group (total speedup effect). Strange though it may seem, *all* researchers in the community save the same time when a collaborative group succeeds more quickly. When the last step is reached, both collaborators and competitors stop searching and unsuccessful researchers lose less time working for nothing. Therefore, every researcher sees an increase in individual successfulness which is proportional to her previous individual successfulness and to the time saved (we call this the individual speedup effect)[9]. In the present case, all researchers have the same individual successfulness in the first configuration, so the total speed up is equally shared between them. In particular, the members of the 2-group each benefit from an increase of $0.296/3 \approx 0.099$. Yet, their increase in individual successfulness is 0.303 (that is, 0.204 more). This corresponds to half of what the lone scientist has lost, namely $0.417 + 0.099 - 0.107 = 0.409$ (which is indeed 2×0.204, up to rounding error). This illustrates the fact that collaborative researchers compete more efficiently and increase their individual successfulness by "stealing" rewards from other competitors. We call this competitive mechanism the *predatory effect*. It is *internal* to the community—some groups steal rewards from others—and it does not affect the total successfulness of the community. In brief, the increase in total successfulness, due to collaboration, comes from the total speedup effect (alone); the increase in individual successfulness, due to collaboration, is accounted for by the individual speedup effect and,

higher total successfulness is.

[8]Our result about individual successfulness can be seen as describing the incentives which the scientists in the models are confronted with (even if they do not know this pay-off structure). The framework of cooperative game theory can then be seen as a promising tool to investigate further this incentive structure, which we leave for future work.

[9]Consider n researchers and the switch from configuration A to B. Researcher i, who used to get individual reward R_i in time t in configuration A, now gets reward $R_i + \delta_{R_i}$ in time $t + \delta_t$ in configuration B (since the time spent is the same for all, there is no index i on t or δ_t). Her individual successfulness is not R_i/t anymore, but $\frac{R_i+\delta_{R_i}}{t+\delta_t} \approx \frac{R_i}{t}(1 - \frac{\delta_t}{t} + \frac{\delta_{R_i}}{R_i})$ at the first order in δ_t and δ_{R_i}. So, at the first order in δ_t, the difference in individual successfulness is proportional to the previous individual successfulness in A. Therefore, it is (at least approximately) meaningful to define the individual speedup effect in such a way that it is proportional to the previous individual successfulness and that all individual speedups sum up to the total speedup.

if there are other competitors, by the predatory effect.

In order to compare the speedup and predatory effects, we plot in fig. 5 the increase in individual successfulness due to each effect for the shifts from configurations (1-1-...-1), in which no one collaborates, i) to configurations (2-1-...-1), in which only two researchers collaborate and ii) to configurations ($(n-1)$-1), in which all researchers but one collaborate. We choose these configurations because both effects are present therein and because they illustrate the possible range of (variations of) degrees of collaboration. We also plot the value of the maximum speedup effect (when shifting from no to full collaboration). The

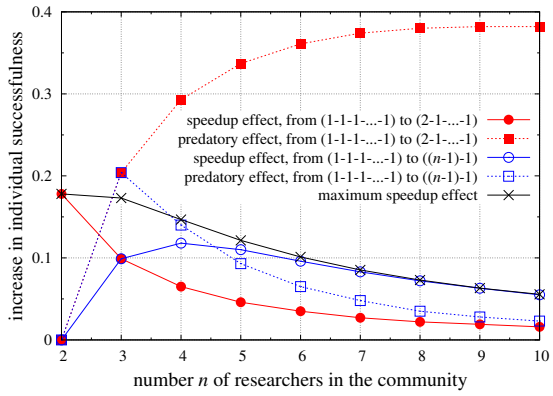


Figure 5: Speedup and predatory effects (for two extreme types of configuration shifts) and maximum speedup effect ($l = 10$ and $p = 0.5$ in all cases).

graph shows that in this model the matter of which effect is larger is a contextual matter. In case i (in red), the predatory effect can be about 20 times larger than the speedup effect, whereas in case ii (in blue), the speedup effect can be two times larger. Further, the maximum speedup curve is much lower than the predatory one in case i. This means that, in this model, when collaboration is massively beneficial, this cannot be due to a significant increase in efficiency but is instead due to a competition effect which amplifies small differential advantages arising from small (sub-additive) increases in group efficiency.

Let us finally analyze the marginal returns of the successfulness functions. The red curve in fig. 3b decreases with n, which means that individuals in fully collaborative groups have no interest in the growth of their group. This is not surprising since $p_g(k, p)$ (the step efficiency of k-groups) grows like $1-(1-p)^k$, that is to say sub-additively. In other words, the total successfulness of fully collaborative groups has decreasing returns, which means that each additional worker hour invested in a program "increases the probability of success a little bit less than the last" (Strevens 2003, 63). This is coherent with the literature on groups, in which marginal returns are usually assumed to be decreasing (Kitcher 1990, 12; Strevens 2003, 63).

However, this feature does not generalize to non-fully collaborative contexts. At first sight, decreasing returns might also be expected, because of the sub-additive growth of $p_g(k, p)$. However, step efficiency and successfulness are two different things, and returns in group successfulness (defined as the sum of the individual successfulness of its k members) can be decreasing *or* increasing, as shown in fig. 6. The solution to this puzzle lies again in the predatory effect. A small increase in step efficiency brought about by a newcomer

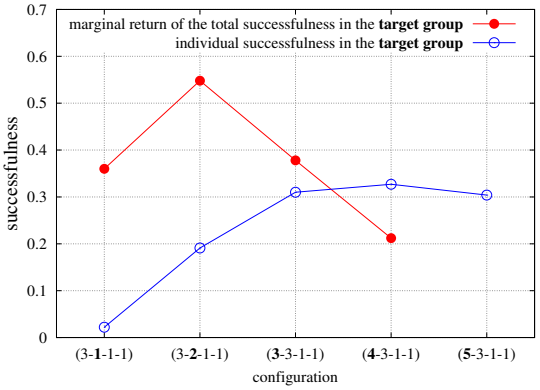


Figure 6: Individual successfulness and marginal return of the total successfulness of a target group (denoted by bold characters) joined by successive newcomers.

can yield massive stealing of the gains of competitors, and, therefore, increasing returns. More precisely, as fig. 6 shows, individual successfulness and marginal return (of group successfulness) behave differently and choosing one quantity or the other would not be equivalent. These variables are however related, since the individual successfulness curve is driven by the difference between the marginal return (if a newcomer was included) and the current individual successfulness. When this difference is positive (whether marginal returns are decreasing or increasing), newcomers increase the individual successfulness of the members of the group by bringing their own share *plus a supplement*. This is a case of super-additive successfulness in which k heads together really are more than k times *more successful* than one. In our model, this simply comes from our conservative hypotheses about collaboration, not from some super-additive hypothesis about the task efficiency of groups. However, as soon as there is nothing left to be stolen, marginal returns collapse, become inferior to individual successfulness (for $k = 4$ in fig. 6), and no longer make up for the inclusion of newcomers. In brief, for a given fixed competitive context, developing collaboration is beneficial up to some point and from a modeling point of view, assuming decreasing marginal returns from the start would distract from some important mechanism which makes collaboration beneficial[10].

5 Drawing some general conclusions.

In spite, or because, of its simplicity, our model can be used to draw general conclusions about collaborating groups, which we present in this section. In the next section, we shall emphasize that these results are not tied to the particular hypotheses of our model and argue for their robustness.

[10]We do not mean that it is always illegitimate to make assumptions about marginal returns or to use this quantity in reward schemes. Kitcher's (1990) and Strevens's (2003) use of marginal returns for retribution is made in a different philosophical context in which one analyzes how researchers should distribute when different programs or methods are competing. We focus on cases with only one research agenda (the sequence of steps) which can be completed in the long run.

5.1 Total successfulness depends on the collaboration structure. Our results show that, in some cases, there is no such thing as *the* total successfulness of a community of n researchers: this quantity depends on how researchers function together in collaborative groups. For example, for $n = 10$ (with $l = 10$ and $p = 0.5$), the total successfulness is 1.998 when researchers fully collaborate and only 1.442 if they do not collaborate at all, which is less than the 1.536 obtained by a community of 2 collaborating researchers! Therefore, for the purpose of the analysis of collaborative groups, it cannot be assumed without careful justification that the prospects of a research community working on a specific problem with a particular method can simply be described as a function of the number of researchers working on it. This is why developing models close to Kitcher's (1990) or Strevens's (2003) would be here somewhat problematic—which points would correspond to their community's successfulness in fig. 3a: the collaborating red ones, the competing blue ones or some in-between purple ones?

5.2 Group successfulness depends on the collaboration structure. There is no such thing as *the* individual successfulness within a k-group, or *the* total successfulness of this group. These quantities depend on the number $n - k$ of other researchers working on the same problem and on their collaboration configuration. For example, a 2-group has a total successfulness of 1.198 in configuration (2-1-1-1-1-1) but only of 0.122 in configuration (5-2), if all single scientists now collaborate. More generally, for a given k, the threshold of $p_g(k, p)$ below which collaboration is no longer beneficial varies from one configuration to another. So, it is not clear exactly what the worst-case analysis of collaboration would be.

5.3 Differences in the step efficiency of groups can result in much larger differences in successfulness. For example, while $p_g(4, 0.5)$ (the probability of a 4-group to pass a step when $p = 0.5$) is 0.9375, $p_g(5, 0.5)$ reaches 0.96875, which is only 3% more. Yet, in configuration (4-5), this leads to a relative difference of 25% in individual successfulness (0.192 vs. 0.242) and 57% in group successfulness (0.768 vs. 1.210). Larger differences in step efficiency can lead to huge differences in successfulness but the extent of this depends greatly on the competition context. For example, $p_g(2, 0.5)$ is equal to 0.75, which is 50% more than $p_g(1, 0.5)$. In configuration (2-1), this leads to a relative difference of almost 700% in individual successfulness (0.720 vs. 0.107) and 1300% in group successfulness.

5.4 Which success mechanisms prevail is a contextual matter. Depending on the collaboration configurations, the predatory effect or the speedup effect can prevail in our model, even if large gains mainly arise from the former. However, other mechanisms may increase the epistemic efficiency of collaborative groups much more than the one we analyze here. Furthermore, in some cases, there may be little gain to steal from other competitors. So, the general conclusion we may draw is that which mechanism (predatory or not) prevails is likely to be a contextual matter.

5.5 The beneficialness of collaboration. Some general conclusions can be drawn about the beneficialness of collaboration although prudence is necessary in doing so. When step-efficiency is super-additive and two collaborators always do more than twice what single scientists do (i.e. not in our model), collaboration is clearly beneficial. When sub-additive

speedup effects improve the step efficiency of the group, as in our model, collaboration can still be beneficial, up to a certain point, as a competition effect, when tiny increases in step efficiency are sufficient to steal the gains of other groups. In any case, when tiny differences in step efficiency are amplified by the competitive regime, it is very unproductive to work alone or in a too small group. In other words, as soon as a community starts collaborating, things become more difficult for lone researchers.

6 Discussing the robustness of the model.

We shall now provide specific evidence that our results are robust under various changes to the model. As we shall see, this robustness analysis is also a way to answer the potential concern that, because our model of collaboration is simple and severely idealized, our results might have limited scope and not be informational regarding more realistic situations.

Firstly, the features discussed in section 5 are not tied to the simplistic collaborative mechanism present in our model. Mathematically speaking, collaboration in our model merely determines the step efficiency of the group. Thus, our 138 data points can be seen as the results of a race between competing (individual or collective) agents with different step efficiencies, whatever the origin of these step efficiencies. Any model with different collaborative mechanisms but which yields identical step efficiencies for these k competitors, gives identical results and conclusions to be drawn provided the problem is sequential and the priority rule applies. The size of groups in these analogous situations need not even be the same: an analogous (3-2) configuration will yield the same group successefulness as our (4-1) configuration if the 3-group (respectively 2-group) has the same step efficiency as our 4-group (respectively 1-group). Beyond exact numbers, this suggests that the features described above are not peculiarities of our conservative model but can be met for other collaborative mechanisms and competing groups which are potentially more realistic than ours.

Secondly, in actual situations, transferring intermediate results between collaborators implies an epistemic cost. However, this is not a threat to our results. In our model, communication takes place when intermediate results are found, that is, at most l times (for winning teams). And the time to take cognizance of and verify the solution of a problem is usually shorter in practice than the time needed to solve it. So if sensible costs for the transfer of information are included in the model, they cannot be large. Finally, communication costs impact both losing and winning teams. This means that winning groups will be slightly less successful but because the above patterns about the successfulness of collaboration are based on large success differences, this will not significantly alter the results.

Thirdly, it is true that our model represents a rather strange and inefficient (and yet possible!) way of collaborating. In general, collaborative groups have deeper and more fruitful interactions. Several cases can be distinguished. These additional interactions may imply an increase in step efficiency. For example, a 2-group may become as efficient as, say, a 3-group in our model. Then, collaboration will be more beneficial, things will be even more difficult for groups which are too small and results about the context-dependence of successfulness will remain. Now, if groups are slightly less efficient (for instance if

the probabilities to pass a step are not perfectly independent anymore for researchers in a group), because collaboration is very beneficial in our model, it will still be, even if somewhat less. Finally, collaborative groups might be *very* inefficient—we do not deny this possibility, even if our model already assumes an inefficient way of collaborating. However, because of their serious unsuccessfulness, such groups are unlikely to be viable and significantly contribute to the observed rise of collaboration in science. Overall, this shows that sensible collaboration is in general beneficial (up to some limit) in a competitive regime. This conclusion could be reached owing to our conservative assumptions about the efficiency of collaboration.

Are changes in the details of the model also inconsequential for our results? Providing evidence in favor of the robustness of our results when p varies from step to step is not straightforward. However, a result by Boyer (2014) about the robustness of a very similar model suggests that, at least when this variation is not too strong, collaborating remains individually preferable[11].

Furthermore, what if all researchers are not equally qualified and p is not constant over individuals? Relevant results can be obtained from our data about equally qualified agents. As pointed out in section 5, our model can be seen as describing a race between competitors with different step efficiencies, whatever the origin of the value of these step efficiencies. Then, by reinterpreting what competitors stand for, we obtain the corresponding data. Suppose some agents in the 3-group in (4-3-1) are more (respectively less) efficient. Configurations (4-4-1) or (5-4-1) (respectively (4-2-1) or (4-1-1)), provide evidence for what happens, since the efficiency of a 3-group is then higher (respectively lower) and may reach that of a 4 or 5-group (respectively 2 or 1-group) or a case in-between. These data confirm the robustness of our morals for such changes.

A similar trick can be used to get evidence about what takes place when p changes for all agents. Since the step-efficiency for a k-group is $p_g(k,p) = 1-(1-p)^k$, groups of 2 (respectively 3, 4 or 5) members with $p = 0.5$ are as efficient as individual researchers with $p = 0.75$ (respectively 0.875, 0.9375 or 0.96875). The data for $p = 0.75$ (respectively 0.875, 0.9375 or 0.96875) are then obtained from the data for $p = 0.5$ by picking configurations in which group sizes are multiples of 2 (respectively 3, 4 or 5), given that, when lumping k individuals together, research time is divided (and individual successfulness multiplied) by k. One thereby obtains exact results for various other configurations up to $n = 5$, for free. The patterns described in the previous section can again be found in these new data. The difference is that, for high p, increases in p soften these features. Typically, in configuration (2-1), the 2-group does about 6.5 times better than the 1-group for $p = 0.5$ (0.720, 0.107), about 4 times better for $p = 0.75$ (0.848, 0.200) and about 2.5 times better for $p = 0.875$ (0.814, 0.352). This is due to the fact that, for $p > 0.5$, the differential advantage in step efficiency for k-groups over individuals (that is $\delta p(p,k) = p_g(k,p) - p = 1-(1-p)^k - p$) is all the less important as p is high.

Finally, some may suspect that our model merely describes cases in which all researchers work on the same problem and belong to the same disciplinary field—whereas collaborative

[11] Theorem 4 from Boyer (2014) can be adapted to the present analogous model to state the following: in the simple case of 2 agents working on a 2-step project with different difficulties p_1 and p_2, it is individually preferable to collaborate if and only if $p_2 < \frac{2-p_1}{1-p_1}p_1$. If research is rather hard, then p_1 is close to 0 and the condition approximates to $p_2 < 2p_1$. So the result states that collaboration remains preferable provided that the difficulty of the second step does not change by more than a factor 2.

science often involves scholars from different fields. We shall now show how to answer this objection. In our model, the various steps of the task may require different types of skills or knowledge and the goal pursued be a multidisciplinary achievement. Therefore, dividing labor by using different experts for each type of step can be seen as a way of increasing the step efficiency of the group. Indeed, when discussing k-group, k denotes the number of agents available on each step of the project. If different experts are in charge with different steps, the workforce (in terms of researchers per step) remains the same but is inferior to the number of individuals involved in the group. Since small differences in step efficiency can imply large differences in successfulness, one can expect collaborations based on division of labor to be extremely successful. Analyzing in detail the effects of the division of scientific labor falls out of the scope of the present article. Nevertheless, our data already show that dividing labor can be an extremely powerful collaborative mechanism. Suppose that the l steps of a problem are ascribed to l specialists (one for each step) with probability 0.75 of passing each step (versus 0.5 for non-specialists). Then, with the same workforce of one researcher for each step, a 1-group of these specialists (corresponding to l cosigning authors and counting as a collaborative group of size l for researchers in scientometrics) is as efficient as a 2-group of non-specialists. When competing with a single non specialized individual, this specialized 1-group gets as much as a non-specialized 2-group in (2-1) (see fig. 4) but with twice less workforce per step. Therefore, instead of getting 0.590 as non-specialized individuals in a (1-1), its members get 1.440 and their non-specialized opponent 0.107. True, it is already known that the division of labor is an efficient mechanism. The current analysis suggests that, under a priority rule, dividing labor can be a devastating weapon and it also shows that our idealized model can be informative about more complex collaborative practices. We leave for future work more precise investigations carried out with a disidealized version of the present model.

Let us finally discuss the maximal number of agents in our simulations. Because the number of configurations grows very quickly with n, we have not gone beyond $n = 10$, which might be deemed small in comparison with the size of actual scientific communities. However, since the parameter n in our model represents the number of researchers working *on the very same scientific problem*, this number is compatible with a research community which is larger than n by several orders of magnitudes. Furthermore, since our model merely includes the sharing of intermediate results between agents, it can be applied to agents who are not individuals but constituted research teams or labs (the question being whether these labs should collaborate or not); and then, 10 labs is already a significant number.

7 Consequences for explanations of collaboration.

The results of our model make it clear that, for individuals, being in a group is better than being alone and large groups usually do better up to a certain extent (cf. section 4). In light of this, can we claim to have provided an explanation of scientific collaboration? Great care should be taken in answering. Showing that individual successfulness is larger when agents collaborate is a valuable point but many philosophers will be reluctant to accept this as an explanation of collaboration. Economists might take it to be obvious that reward structures affect behavior, but philosophers of the social sciences with inclinations

for causal explanations may require it to be shown how the beneficialness of collaboration causes the existence and persistence of collaborative behavior (Kincaid 1996). It is true that connections between the reward structure presented here and the existence of actual collaboration patterns are plausible. For example, scientists may use clues to make approximate but partly reliable judgments about when and how much it is appropriate to collaborate. Collaborating scientists may become successful and develop prosperous research environments, so that collaborating becomes a norm which is transmitted to more and more researchers (Wray 2002). Since defending any such explanations would require careful work to be done convincingly and this explanatory goal is additional to and partly independent from the achievements completed so far, we leave this for future work and acknowledge that our present results do not by themselves provide an explanation of collaboration.

Nevertheless, we now argue that our model provides important insights regarding what such explanations can or should look like.

Firstly, our sufficiency results about the potential effects of intermediate results sharing are important in a modeling explanatory perspective. The reason is that, while there may be a great variety of factors which improve the step efficiency of groups, sharing intermediate results is a basic feature of collaborative practices, for problems that can be modeled as sequential. Therefore, its effects are also at work when other mechanisms are present. In any case, methodologically speaking, to analyze complex cases and assess the relative weigh of other beneficial factors (such as cross-checking or the generation of collective ideas), one needs to estimate the potential effects of basic and ubiquitous features such as intermediate result sharing. Further, since this simple mechanism is a general feature of collaborative groups, which is sufficient to make collaboration beneficial, anyone who might like to idealize it away and give a central role to other mechanisms needs to show why or when the effects of these mechanisms supersede the effects of intermediate results sharing. Our analyses in section 6 suggest that situations in which labor can be divided between specialized experts is probably such a case.

Secondly, our results highlight important constraints for EB_E-accounts, which explain the epistemic efficiency of groups in terms of epistemic factors only. Since the actual impact of the epistemic properties of agents can be context-sensitive and emerge when these agents compete together (see section 4), group successfulness cannot be intuitively inferred from the epistemic properties of individuals. Instead, unexpected features, like bare intermediate results sharing, may play an explanatory role. Thus, one should be especially wary of non-contextual EB_E accounts. Anyone defending such accounts would need to show that the epistemic effects of the epistemic features of groups do not significantly depend on the larger context in which these groups work. For cases in which predatory gains may exceed those due to increases in epistemic efficiency (see section 4), EB_E accounts are even more difficult to defend. For instance, let us suppose that collaboration makes k-groups work q times quicker, with $q > k$, whatever the context and also that, in some contexts, the successfulness of k-groups is actually multiplied by r *because of the predatory effect*, with $r \gg q$. Then, the purely epistemic gain of q may sometimes be insufficient to account for the existence of a collaborative activity (and the r gain sufficient).

Let us now present some positive conclusions that can be drawn. First, EB-accounts of collaboration which are local in the sense that they emphasize the benefits of collaboration

at the step level, have more powerful versions in which these local benefits are amplified by the priority rule. In particular, EB_E accounts but also partly social EB_{E+NE} accounts, like those emphasizing the benefits of specialization and the division of labor, can be turned into stronger accounts when the conditions of the models are met.

Second, even minor local effects, coming from epistemic or non-epistemic factors at the step level, can have an important impact on the final successfulness of a research group *at the research project level*, a larger impact than would have been expected without a quantitative model. Therefore, strong collaborative mechanisms are not necessary to explain collaboration; further, even in the presence of strong factors, minor factors can still be sufficient to make a difference between teams.

8 Conclusion

In a competitive environment, increases in group successfulness can come from epistemic gains in efficiency (speedup effect) and gains that are made at the expense of other competitors (predatory effect). Even if the matter which effect prevails is contextual, our results show that the predatory effect can be huge and single scientists or too small groups can be hardly viable in a collaborating community. Also, small differences in epistemic step efficiency, like those coming from the sharing of intermediate results in the case of sequential tasks, can lead to massive differences in final successfulness. This is why two heads need not be more than twice better than one (even if they probably often are) for collaboration to be beneficial.

We have also shown that, when the priority rule applies, the success of groups and communities depends on the collaboration structure of the community. Even if our results do not by themselves provide a fully-fledged account of collaboration, we have emphasized that they act as restrictions on explanations of collaborations and make it difficult to draw explanatory inferences about the success of collaboration from the epistemic properties of groups *simpliciter*. Finally, because of the possible amplification of small differences in efficiency, minor epistemic factors can sometimes play a decisive role and potentially strong explanatory factors like the division of scientific labor and specialization can be even more powerful, which overall suggests that socio-epistemic explanations of collaboration are more plausible.

Acknowledgments

Earlier versions of this paper were presented at the workshops " Epistemic groups and collaborative research in science" and "Modeling epistemic and scientific groups: interdisciplinary perspectives" in Nancy (France). We would like to thank the discussants for their helpful comments. We also thank the referees for *Philosophy of Science* for their insightful reports. This work was supported by the MSH Lorraine (project "Colexia"), the Archives H. Poincaré, the Université de Lorraine (Nancy, France), and by STL (CNRS and Université de Lille 1 and Lille 3, France).

References

Anderson, Katharine A. (2013), "The Formation of Collaboration Networks Among Individuals with Heterogeneous Skills", GSIA Working Papers, unpublished.

Batterman, Robert W. (2002), "Asymptotics and the Role of Minimal Models", *The British Journal for the Philosophy of Science* 53 (1): 21–38.

Beaver, Donald deB. and Rosen, Richard (1978), "Studies in Scientific Collaboration: Part I", *Scientometrics* 1(2): 65–84.

— (1979a), "Studies in Scientific Collaboration: Part II", *Scientometrics* 1(2): 133–149.

— (1979b), "Studies in Scientific Collaboration: Part III", *Scientometrics* 1(3): 231–245.

Boyer, Thomas (2014), "Is a bird in the hand worth two in the bush? Or, whether scientists should publish intermediate results", *Synthese* 191: 17–35.

De Langhe, Rogier (2014), "A unified model of the division of cognitive labor", *Philosophy of Science* 81(3): 444–459.

Fallis, Don (2006), "The Epistemic Costs and Benefits of Collaboration", *Southern Journal of Philosophy* 44 S: 197–208.

Goldman, Alvin (1999), *Knowledge in a Social World*, Oxford University Press.

Hardwig, John (1985), "Epistemic Dependence", *The Journal of Philosophy* 82(7): 335–349.

Kincaid, Harold (1996), *Philosophical Foundations of the Social Sciences: Analyzing Controversies in Social Research*, Cambridge University Press.

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

Latour, Bruno (1987), *Science in action: How to follow scientists and engineers through society*, Cambridge, MA: Harvard University Press.

Muldoon, Ryan (forthcoming), "Diversity, Rationality and the Division of Cognitive Labor", in Boyer-Kassem, T., Mayo-Wilson, C. and Weisberg, M. (eds.), *Scientific Collaboration and Collective Knowledge*, New York: Oxford University Press.

Muldoon, Ryan and Weisberg, Michael (2011), "Robustness and idealization in models of cognitive labor", *Synthese* 183(2), 161–174.

Rennie, Drummond, Flanagin, Annette and Yank, Veronica (2000), "The contributions of authors", *Journal of the American Medical Association* 284 (1), 89–91.

Solomon, Miriam (1994), "Social Empiricism", *Noûs* 28 (3): 325–343.

Schelling, Thomas (1971), "Dynamic Models of Segregation", *Journal of Mathematical Sociology*, 1:143–186.

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

Thagard, Paul (2006), "How to Collaborate: Procedural Knowledge in the Cooperative Development of Science", *The Southern Journal of Philosophy* XLIV: 177–196.

Weisberg, Michael (2007), "Three Kinds of Idealization", *The Journal of Philosophy* 104(12): 639–59.

Wray, K. Brad (2002), "The Epistemic Significance of Collaborative Research", *Philosophy of Science* 69 (1): 150–168.

— (2006), "Scientific authorship in the age of collaborative research", *Stud. Hist. Phil. Sci.* 37: 505–514.

Wuchty, Stefan, Jones, Benjamin F. and Uzzi, Brian (2007), "The Increasing Dominance of Teams in Production of Knowledge", *Science* 316(5827): 1036–1039.

Zollman, Kevin J. S. (2007), "The Communication Structure of Epistemic Communities", *Philosophy of Science* 74: 574–587.

— (2010), "The Epistemic Benefit of Transient Diversity", *Erkenntnis* 72: 17–35.