The Power of Diversity over Large Solution Spaces

Marco LiCalzi, Oktay Surucu,

The Power of Diversity over Large Solution Spaces

Marco LiCalzi

Department of Management, Università Ca' Foscari Venezia, 30121 Venezia, Italy,
licalzi@unive.it

Oktay Surucu

Department of Economics and Finance, Università LUISS Guido Carli, 00198 Roma, Italy,
osurucu@luiss.it

We consider a team of agents with limited problem-solving ability facing a disjunctive task over a large solution space. We provide sufficient conditions for the following four statements. First, two heads are better than one: a team of two agents will solve the problem even if neither agent alone would be able to. Second, teaming up does not guarantee success: if the agents are not sufficiently creative, even a team of arbitrary size may fail to solve the problem. Third, *defendit numerus*: when the agent's problem-solving ability is adversely affected by the complexity of the solution space, the solution of the problem requires only a mild increase in the size of the team. Fourth, groupthink impairs the power of diversity: if agents' abilities are positively correlated, a larger team is necessary to solve the problem.

1. Introduction

Diversity is a word loaded with multiple connotations. We inquire about the value of bringing a diverse toolbox to a difficult problem. This is well captured in an amusing story by Feynman, a famous and eccentric physicist:

> It turns out that [differentiation under the integral sign] is not taught very much in the universities; they don't emphasize it. But I caught on how to use that method, and I used that one damn tool again and again...The result was, when guys at MIT or Princeton had trouble doing a certain integral, it was because they couldn't do it with the standard methods they had learned in school...Then I come along, and try differentiating under the integral sign, and often it worked. So I got a great reputation for doing integrals, only because my box of tools was different from everybody else's.... (Feynman and Leighton 1985, pp. 49–50)

Taking a managerial point of view, it is natural to expect that a team working on a problem benefits both from the sheer ability of its members and from the diversity of their toolboxes. For instance, upon being honored for excellence in diversity, Dalip Raheja, president and chief executive officer of the Mpower Group, states explicitly that the added value of this consulting firm is in providing diversity of thought: "We want the client to have the benefit of getting all different types of thinking applied to their problem" (Kornik 2008).

More explicitly, here is how Paul Bürkner, president and chief executive officer of the Boston Consulting Group, describes the power of diversity in problem solving:

> When we're faced with what looks at first like an unsolvable problem, a team with what I call 'spikes' of different talents will come up with a better solution than a team whose members have similar strengths. One person makes an oddball suggestion, the next person misunderstands it but in a fruitful way, and together they end up devising a novel solution. The process can be slow and uncomfortable; spikiness often hurts. But it can yield spectacular results. (Dowling 2007)

The power of diversity in problem solving, recently popularized in Page (2007), has been long recognized by the psychological literature (Michaelsen et al. 1989). The consensus over the available evidence distinguishes between *intellective* and judgmental tasks (Hill 1982). The former requires the group to achieve a correct answer, and the latter strives for consensus (Laughlin 1999). It is apparent that groups tend to perform better than individuals on intellective tasks (Hastie 1986, Levine and Moreland 1998). This effect is particularly evident when, as in Feynman's feats, the intellective task is *demonstrable* and the correct solution (once found) is easily recognizable by the other members of the team (Davis 1992, Laughlin et al. 2002).

Some recent literature in experimental economics focuses on intellective tasks. Blinder and Morgan (2005) observe that many important decisions, most notably in monetary policy, are made by committees rather than individuals. Comparing two laboratory experiments based on a demonstrable intellective task, they reach three conclusions: groups make better decisions than individuals; they require less information to act; and the quality of their decision is not affected by whether decisions are made by majority or unanimity. In a follow-up study, Blinder and Morgan (2008) provide evidence that designating the ablest individual as team leader has no discernible effects on performance. A different strand has worked on game-theoretic intellective tasks, reaching the general conclusion that initially teams may not perform better than individuals but they learn faster to adopt and use strategic reasoning, from which they gain a substantial payoff advantage; see the beauty contest in Kocher and Sutter (2005), the limit pricing game in Cooper and Kagel (2005), and the centipede game in Bornstein et al. (2004).

More generally, the literature on organizational and management science shares the insight that functional diversity with a group tends to improve its performance (Krishnan et al. 1997, Simons et al. 1999, Jackson et al. 2003). At the same time, much effort has gone into factoring out a variety of details that promote or hinder the performance of a team (Pelled et al. 1999, Mannix and Neale 2005). For instance, Dahlin et al. (2005) argue that educational diversity favors information use although the social categorizations cued by national differences impair it. Reagans et al. (2004) show that the heterogeneity of the social networks associated with its members is an important predictor of team performance.

Finally, even the ecological literature has recently turned its attention to experiments providing evidence that groups may perform better than individuals; see Conradt and List (2009) for a wide-ranging survey. Liker and Bókony (2009) show that larger groups of house sparrows exhibit a faster and greater foraging performance; their superiority is not only due to a higher number of attempts but to a greater effectiveness in problem solving. Burns and Dyer (2008) argue for the evolutionary advantage of maintaining a diverse set of foraging strategies in a bumblebee colony. Ward et al. (2011) point out that both speed and accuracy of decision making increase with group size in fish shoals under predation threat.

The aim of this paper is to study the impact of diversity over individual problem-solving abilities when a team faces a demonstrable intellective task over a large solution space. We assume that the solution must be found in a very large solution space and that agents' abilities are constrained to explore only limited regions of such space. We exhibit sufficient conditions under which a team of two or more agents succeeds in solving a difficult problem with a large solution space: even though no agent is sure to crack it, pooling their diverse abilities achieves the goal. In other words, we prove when it takes (at least) two to solve a problem. We also provide sufficient conditions under which teaming up does not always guarantee success because agents are not sufficiently creative.

Moreover, we also consider situations in which the problem-solving ability of agents is adversely affected by the dimension of the solution space. We prove how a small increase in the size of the group can be highly effective to overcome the negative effects on agents' abilities. A second negative effect is associated with groupthink, when the members of a team tend to align their perspectives. We show that increasing groupthink induces progressively greater losses in the group performance, but that these can to a large extent be compensated by substantially stronger increases in team size.

The organization of this paper is as follows. Section 2 describes the main model, accompanied by an illustrative example. Section 3 shows that a team of at least two agents may be always successful whereas a single agent would not. Section 4 discusses the equivalence of our model with the one used in Hong and Page (2004), highlighting the differences in our perspective. Section 5 considers a model where teaming up may fail to ensure success, regardless of the size of the team. Section 6 deals with the assumption that problem-solving abilities are inversely related to the size of the solution space and §7 with the effects of groupthink. Section 8 offers a commentary. Proofs are relegated to the appendix.

2. The Model, Illustrated by an Example

There is a team T of m problem-solving agents of limited ability who attempt to maximize an objective function V that maps a finite set X of n solutions into real numbers. The function $V: X \to \mathbb{R}$ is one-to-one; in particular, it has a unique maximizer at x^*. The task of the team is locating x^*. Following Marschak and Radner (1972), we assume that the agents in the team share the same objective and there are no frictions due to the difficulties of communication among people.

As a simple running example, consider Figure 1: It depicts a solution space X with six elements. We identify each point x with its value $V(x)$ topping it; hence, the maximizer $x^* = 6$ is the bottom-left element associated with a value of 6.

Locating the maximizer is a demonstrable intellective task that can be carried out *disjunctively*: If one of the agents finds x^*, he can show its value to his

Figure 1   A Solution Space with Six Elements
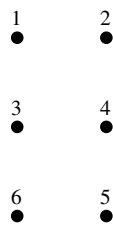
teammates and the task is accomplished (Gerchak and He 2004). However, because of their limited ability, each of the agents in the team may fail to discover x^*.

We represent the limited problem-solving ability of each agent i by a partition Π_i of X; see Rubinstein (1993) for a related approach. We momentarily suppress subscripts and consider one agent. The mutually disjoint and exhaustive classes constituting the partition Π are called *blocks*. The agent can find the best solution within the block he is working on, but he is impervious to the other blocks unless his attention is redirected there by someone else. The *availability map* $\Pi: X \rightarrow 2^X \setminus \varnothing$ describes the search space of the agent: for each $x \in X$, $\Pi(x)$ is the set of solutions that he can explore when he is aware of a candidate solution x. The availability map is *consistent*: x is in $\Pi(x)$ for all x in X; that is, the search space of an agent always contains the candidate solution.

We assume that each agent correctly identifies and compares the values of V for each solution he examines. Therefore, given a candidate solution x, he explores the search space $\Pi(x) \subseteq X$ and finds $x_0 = \arg\max_{y \in \Pi(x)} V(y)$: Because $V(x_0) \geq V(x)$, the search is always (weakly) improving. More generally, when the agent has access to an initial subset $S \subseteq X$ of candidate solutions, he explores $\Pi(S) = \bigcup_{x \in S} \Pi(x)$ and finds the solution $x_0 = \arg\max_{y \in \Pi(S)} V(y)$.

Continuing our running example, Figure 2 depicts a partition for a (male) agent who thinks "horizontally": Working over the top (respectively, middle or bottom) block of possible solutions, the best solution he can find is 2 (respectively, 4 or 6). However, by himself, the agent lacks the ability to move from one block to another. Unless he happens to begin his search in the bottom block, he cannot find $x^* = 6$.
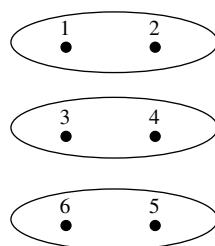
When two or more agents work together, they can pool their abilities and expand their search spaces. Following the general framework developed in Hong and Page (2001), it is not necessary to specify the minute details of their interaction because there are many alternative procedures that lead to the same outcome. However, for the sake of clarity, we describe a specific procedure inspired by sequential search. Number agents from 1 to m. Agent 1 works on the task from an initial set S_0 of candidate solutions. He discovers a set $S_1 = \Pi_1(S_0) \supseteq S_0$ of possible solutions that he hands over to agent 2. She then explores the search space $S_2 = \Pi_2(S_1) \supseteq S_1$. In each stage, the tentative solution is $x_i^* = \arg\max\{V(x): x \in \Pi_i(S_{i-1})\}$; clearly, $V(x_i^*) \geq V(x_{i-1}^*)$ and the search is (weakly) improving. The process continues, cycling over all permutations of the agents, until no one can find further improvements; then the search stops and the tentative solution becomes final. If, along the process, no agent ever gets to explore the block in his partition that contains x^*, the optimal solution is not found and the team fails (although it may succeed in discovering a very good local optimum).

Continuing with our illustrative example, suppose that there is a second (female) agent with the "vertical" partition shown in the center of Figure 3. She teams up with the first agent, whose partition from Figure 2 is reproduced on the left of Figure 3. The male agent starts with $S_0 = \{1\}$ and discovers the set of solutions $S_1 = \Pi_1(S_0) = \{1, 2\}$ in his top block. He hands this information to the female agent, who can explore $S_2 = \Pi_2(S_1) = \{1, 2, 3\}$ and thus finds $x = 3$. Exchanging information, she passes her insight back to her teammate. As $S_3 = \Pi_1(S_2) = \{1, 2, 3, 4\}$, this unlocks his access to $x = 4$. When the male agent returns this new piece of information, this opens her eyes on her bottom-right block. She finds out $x = 5$ and communicates it to the male agent, leading him to finally discover $x^* = 6$.

We formalize the problem-solving ability of a team $T = \{1, 2, \ldots, m\}$ as follows. Each agent i in T is associated with a partition Π_i that represents his ability to explore the search space. Starting with an initial set S of candidate solutions, each agent i explores $\Pi_i(S)$; analogously, any team of two agents $i = 1, 2$ explore $\Pi_i(\Pi_{3-i}(S))$; and, more generally, the team

Figure 2   The Partition of an Agent

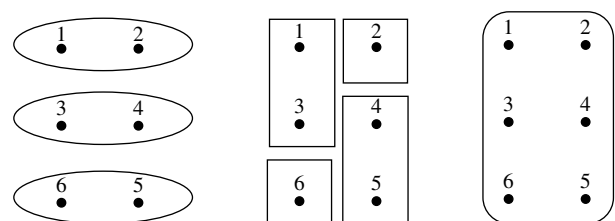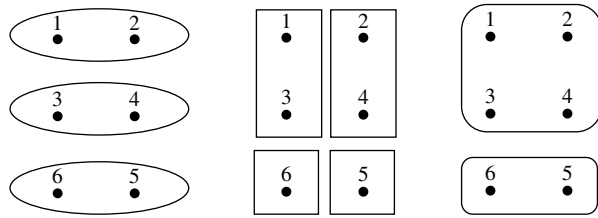**Figure 3   The Partitions of Two Agents and Their Meet**

Figure 4   The Partitions of Two Other Agents and Their Meet

jointly explores the union of all subsets representable as $\Pi_{i_k}(\Pi_{i_{k-1}} \ldots (\Pi_{i_1}(s))) \subseteq S$ for any k and any sequence $i_1, i_2, \ldots, i_k$.

When X is finite, there is a simple characterization of the set of solutions that are jointly explored by a team. The finest common coarsening of the partitions $(\Pi_i, i \in T)$ is another partition M called their *meet*. Then, from an initial subset S of candidate solutions, the team explores $M(S)$. In other words, the meet describes the problem-solving abilities jointly attained by the team.

For instance, the meet of the partitions of the two agents used in our example is represented on the right of Figure 3. This is the trivial partition: Regardless of its starting point, the team can access the entire space of solutions and thus it is always successful. On the other hand, things may not work as smoothly. For instance, switch her right blocks and suppose that the second agent comes with the partition given at the center of Figure 4. If the two agents work on any subset of $S = \{1, 2, 3, 4\}$, they are jointly unable to access any point $x \geq 5$ and will not find $x^* = 6$. The meet of their partitions is shown on the right of Figure 4: It reveals that the team, albeit more powerful than each of its members alone, still exhibits limited problem-solving abilities.

The next short subsection conveniently collects a few technicalities used in this paper.

2.1. Technicalities

We assume that X is a finite set with n elements; when useful, we write it as X_n to make the number of elements evident. We label the elements of X_n with the integers $\{1, \ldots, n\}$. Following custom, we list the blocks of a partition of X in increasing order of their least elements and the elements of each block in increasing order. For instance, the blocks of the partition $\{3, 4, 5\}, \{6, 1\}, \{2\}$ of a set with six elements are listed $\{1, 6\}, \{2\}, \{3, 4, 5\}$. For brevity, we simplify notation and write the partition as $16 \mid 2 \mid 345$.

We write $\Pi \preceq \Pi'$ to denote that Π is coarser than Π'. Clearly, an agent endowed with partition Π has a higher problem-solving ability than another agent endowed with Π'. The trivial partition Π_0 that has X as its unique block satisfies the property $\Pi_0 \preceq \Pi$ for any partition Π; therefore, an agent endowed with the

trivial partition has the highest problem-solving ability of all and, indeed, will find the global optimum from any starting point. The set of all the partitions of X partially ordered by the refinement relation $\preceq$ is a lattice. In particular, the notation $\Pi_1 \wedge \Pi_2$ denotes the finest coarsening of Π_1 and Π_2; analogously, $\Pi_1 \vee \Pi_2$ stands for their coarsest refinement. The meet of $(\Pi_i, i \in T)$ is $M = \bigwedge_{i \in T} \Pi_i$.

The number of partitions for a (finite) set X_n of n elements is given by the *Bell number* B_n. As an example, consider $X_3 = \{1, 2, 3\}$. The set of all its partitions, denoted by $\mathscr{P}(X_3)$, is $\mathscr{P}(X_3) = \{123, 1|23, 2|13, 3|12, 1|2|3\}$ and thus $B_3 = |\mathscr{P}(X_3)| = 5$. The first few Bell numbers are $B_0 = 1$, $B_1 = 1$, $B_2 = 2$, $B_3 = 5$, $B_4 = 15$, $B_5 = 52$, and $B_6 = 203$. The Bell numbers satisfy the recursive formula:

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k,$$

as well as the Dobinsky's formula:

$$B_n = \sum_{k=0}^{+\infty} k^n \cdot \left(\frac{e^{-1}}{k!} \right),$$

according to which B_n is the nth moment of a Poisson distribution with expected value of 1.

We are to study spaces of solutions when n is large. Clearly, the number B_n of possible partitions is much larger than n. Some intuition for its rate of growth can be gathered considering that $B_{10} = 115{,}975$; $B_{100} \approx 4.7585 \times 10^{115}$; and $B_{1{,}000} \approx 2.9899 \times 10^{1{,}927}$. An asymptotic formula for the Bell numbers as $n \uparrow +\infty$ is

$$B_n \sim \frac{1}{\sqrt{n}} r^{n+1/2} e^{r-n-1},$$

where r is defined as the root of $re^r = n$; see Pitman (2006). Recently, Berend and Tassa (2010) have proved a convenient upper bound,

$$B_n < \left(\frac{0.792n}{\ln(n+1)} \right)^n, \tag{1}$$

that holds for all n.

3.   Two Heads Are Better Than One

This section studies the performance of a randomly chosen team of problem solvers. We fix both the number n of solution in X_n and the size m of the team T_m. A team is *always successful* if it can find the optimal solution in X_n from any starting point; that is, if the meet of the partitions of its agents is the trivial partition. We assume that the agents in the team are randomly chosen so that the composition of the team is stochastic and we ask what is the probability that the team is always successful.

Table 1   When the Meet of Two Partitions Is the Trivial Partition

	No difficulty					Difficulty				
	123	1|23	2|31	3|12	1|2|3		1|23	2|31	3|12	1|2|3
123	×	×	×	×	×					
1|23	×		×	×		1|23		×	×	
2|31	×	×		×		2|31	×		×	
3|12	×	×	×			3|12	×	×		
1|2|3	×					1|2|3				

Because a problem solver is represented by a partition, we need to construct a model of how agents are randomly drawn from the set of partitions of X_n. To stack the deck against success, we borrow from Hong and Page (2004) the assumption that no single agent can find the optimum alone.

ASSUMPTION 1 (DIFFICULTY). *No agent in the team is endowed with the trivial partition.*

The simplest way to build a random model for the problem-solving abilities of the members of a team is to assume that they are independently chosen according to the uniform distribution over all partitions. Formally, let $\mathscr{P}(X_n)$ denote the set of all partitions of X_n. The number of elements of $\mathscr{P}(X_n)$ is the Bell number B_n. The uniform distribution on $\mathscr{P}(X_n)$ assigns probability B_n^{-1} to each partition. We call this the *uniform model*. To satisfy the difficulty assumption, it suffices to attach zero probability to the trivial partition and update the uniform model by giving probability $(B_n - 1)^{-1}$ to every nontrivial partition. We call this the *uniform model with difficulty* and study it in this section. Sections 5 and 6 consider two alternative models.

It is intuitively clear that the size n of the set X_n of solutions is going to affect how easy it is for a team to always be successful. The main focus of this paper is over the case when n is large. However, to build up intuition, we look first at the case where n is quite small.

Consider two agents and three elements, so that $m = 2$ and $n = 3$ with $X = \{1, 2, 3\}$. There are $B_3 = 5$ possible partitions, namely, 123, 1|23, 2|31, 3|12, 1|2|3. Each of these five partitions has identical probability 1/5 and each pair of partitions (one for each agent) has identical probability 1/25 of occurring. Denoting by × the event that the meet of two partitions is the trivial coarse partition, we obtain the left-hand side of Table 1 where it is easy to check that there are 15 favorable events out of 25 possible ones. Because

the joint probability distribution is uniform, the probability that a team of two agents is always successful under the uniform model is $15/25 = 0.6$.

Under the difficulty assumption, the probability that an agent is endowed with the trivial partition 123 is zero, so we delete the first row and the first column and get the right-hand side of Table 1. Now, there are six favorable events out of 16. Thus, under the uniform model with difficulty, the probability that a team of two agents is always successful is $6/16 = 0.375$. Introducing difficulty reduces the probability of success from 0.6 to 0.375; as expected, it is harder to find the optimal solution.

Using a similar reasoning, we have computed the probability that a team of $m = 2$ agents is always successful under the uniform model with difficulty for $n = 2, 4, 5$ as well. The results are summarized in the first row of Table 2, whereas the second row reports the same probabilities without the difficulty assumption.

We make a few observations. The probability that a team is always successful is definitely smaller than 1, regardless of whether or not we assume difficulty. Under the uniform model, a team of two person may fail to solve the problem even if the solution space is very small. This provides an important baseline for appreciating the result we are going to prove below; namely, as the size n of the solution space X_n grows, the probability that a team of two people is always successful approaches 1. That is, for a fixed team size, expanding the solution space makes the team much better at problem solving.

A hint of this surprising result shows up in Table 2 if we look at the first row from left to right. As n increases, the ability of the team of always solving the problem under difficulty increases from 0 to 0.477. What is going on? The basic idea is that a larger solution space makes it easier for new ideas or novel approaches to emerge. Given the current set S_t of candidate solutions, the team can make progress only if the search space $\Pi_i(S_t)$ accessible by some agent i contains some element x that is new with respect to S_t. Because agents can transfer their insights to teammates, such an additional element x expands the team's solution space from S_t to $S_t \cup \{x\}$ and leads to a (weak) improvement. Each fresh insight x may suggest a new approach and open up new possibilities. When the search space grows large, there are more items that can be discovered: each of them can be a

Table 2   Probability That a Team of Two Agents Is Always Successful

	$n = 2$	$n = 3$	$n = 4$	$n = 5$
Difficulty	0	$6/16 = 0.375$	$90/196 = 0.459$	$1{,}240/2{,}601 \approx 0.477$
No difficulty	$3/4 = 0.750$	$15/25 = 0.600$	$119/225 \approx 0.529$	$1{,}343/2{,}704 \approx 0.497$

Table 3 Probability That a Team of Two Agents Is Always Successful (Estimates)

	$n = 25$	$n = 50$	$n = 75$	$n = 100$
No difficulty	0.679	0.797	0.847	0.871

stepping stone or a *bridge* to a better solution. Creativity is fostered by tossing around many alternatives.

One last technical observation compares the values in Table 2 column by column. The difference between the probability that a team is always successful is decreasing in the size n of the solution space X_n. In other words, as the solution space expands, the impact of the difficulty assumption on the ability of the team to solve the problem tends to be negligible. Indeed, it turns out that it is asymptotically irrelevant.

We are ready to consider what is the probability that a team of m people is always successful under the uniform model when the space X_n of solutions grows large and $n \uparrow \infty$. We find that this probability approaches one for any $m \geq 2$, with or without difficulty. On the other hand, the probability that an agent is always successful is zero under the difficulty assumption, whereas it goes down to zero as $1/B_n$ without it. This is a sharp difference: when the space of solutions is sufficiently large, we are almost sure that no single agent is able to solve it but a team of just two (or more) people will. Tossing many solutions around unclenches the power of diversity.

THEOREM 1. *Under the uniform model, the probability that a team of at least $m \geq 2$ people is always successful tends to 1 as n increases.*

For n large, an exact computation of the probability is prohibitive because of the size of the Bell numbers involved. However, Equation (7) in the appendix provides an asymptotic estimate for the rate at which this probability approaches 1. Moreover, as suggested by one referee, we have estimated the probabilities (assuming no difficulty) for a few different values of n over 10,000 rounds of simulations using a plain Monte Carlo approach. These are reported in Table 3.

4. A Comparison with Hong and Page

The seminal reference for the power of diversity is the work of Hong and Page (2001, 2004), recently popularized in Page (2007). This section shows that the model used in Hong and Page (2004) for the limited problem-solving ability of an agent is equivalent to ours and that we abide by the same three main assumptions (Difficulty, Diversity, and Uniqueness) made there. However, our perspective is novel in two important respects.

First, from a methodological point of view, we introduce partitions as an effective tool to represent the limitations in problem-solving ability. Second, from a substantial point of view, the main result

in Hong and Page (2004) assumes that the number m of agents is sufficiently large but the solution space X_n has a fixed size n. We take the opposite point of view and let n grow large while keeping fixed the cardinality of the team of problem solvers who are in charge of tackling it. They are interested in large teams; we study large solution spaces. So we view the two papers as complementary with regard to exploring the power of diversity.

We also argue that proving the power of diversity over a large solution space X_n is more challenging than establishing it for a large team of m agents. If we assemble a larger team to work over a fixed solution space, it stands to reason that it should have a better chance to succeed. On the other hand, consider a team of fixed size working on an ever larger solution space. It is far less obvious that the team should always be successful. A simple example may help to ground this intuition.

Table 4 exhibits the probability that a team of $m = 2, 4, 8, 16$ agents is always successful for $n = 3$ or $n = 4$ under the uniform model with difficulty.

All values are rounded to the closest fifth decimal digit, and we write 1^- when the probability is within 10^{-5} from one. Assuming ever larger teams over a given solution space X_n corresponds to reading a row in the table from left to right: as expected, for small values of n, the probability that a team of m agents is always successful is rapidly approaching one. Our paper, on the other hand, is concerned with what happens when we descend a column in the table.

The general framework in Hong and Page (2001) describes the problem-solving ability of an agent as the pairing of his perspective with a set of heuristics. The *perspective* is the agent's internal representation of a problem; the *heuristics* are the algorithms he applies to locate solutions. To facilitate the mathematical study of disjunctive tasks, Hong and Page (2004) suppress their own distinction between perspectives and heuristics and characterize each agent i by a mapping $\varphi_i: X \rightarrow X$ and a probability distribution ν on X. The initial distribution ν has full support and it is the same for all agents: it is used to randomly generate the starting point of the search process; without loss of generality, let ν be the uniform distribution on X.

For each x, $\varphi_i(x)$ denotes the local solution found by agent i when he starts his search at x. Assumption 0 in Hong and Page (2004) states two properties valid

Table 4 Probability That a Team of m Agents Is Always Successful over X_n

	$m = 2$	$m = 4$	$m = 8$	$m = 16$
$n = 3$	0.37500	0.82031	0.98831	0.99995
$n = 4$	0.45918	0.91977	0.99881	1^-

for any mapping φ_i and encapsulates the hypothesis that each agent i is intelligent:

(0.a) $V(\varphi_i(x)) \geq V(x)$ for all x in X;

(0.b) $\varphi_i(\varphi_i(x)) = \varphi_i(x)$ for all x in X.

Assumption (0.a) is the obvious requirement that the agent never finds a local solution worse than his starting point. Assumption (0.b) states that the local solution found starting at x cannot be further improved upon by the agent and that the final point of his search is unique. This is natural if we take the point of view that agent i tries his best given his set of heuristics.

Under Assumption 0, it takes a simple change of perspective (pun intended) to derive the equivalence between our partitional model and the φ-representation. Suppressing subscripts momentarily, consider an agent associated with the mapping φ and the equivalence relation $x \sim y$ on $X \times X$ defined by $\varphi(x) = \varphi(y)$. When $x \sim y$, the agent starting his search at either point ends up discovering the same local optimum. With respect to the objective of maximizing V, he is indifferent between x and y because they both lead to an identical result. Hence, $\sim$ defines an indifference relation that partitions X into equivalence classes such that each starting point in the same class leads to the same local maximum.

Clearly, the problem-solving ability of an agent represented by φ is uniquely identified with the partition Π induced by φ. In this respect, it is worth noting that Assumption (0.b) implies $x \sim \varphi(x)$ for any x and thus plays the important role of ensuring that φ induces a partition Π that is consistent in the sense defined in §2.

Vice versa, any partition Π of X uniquely defines a problem-solving mapping φ by the following construction. Let $\sim$ be the equivalence relation on $X \times X$ defined by $x \sim y$ if and only if $y \in \Pi(x)$, where $\Pi(x)$ denotes the block of the partition Π that contains x. Then, for each x in X, the mapping $\varphi(x) = \arg\max_{y \sim x} V(y)$ characterizes the problem-solving ability of the agent endowed with the partition Π. This establishes a formal equivalence between the mapping φ and the partition Π as models of limited problem solving for an agent.

This formal equivalence allows us to rephrase results from one perspective to the other. For instance, Hong and Page (2004) note that the image $\varphi(X)$ of the mapping is the set of local optima discoverable by the agent. Because the elements in $\varphi(X)$ are in a one-to-one correspondence with the blocks in the partition Π, the cardinality of $\varphi(X)$ is the same as the number of blocks in Π. More generally, in the Hong and Page (2004) model, an agent is more creative if he has more heuristics (or possibly more perspectives) and thus can generate more points in the search space; in

our model, this naturally translates into having a partition made of larger blocks.

In a similar vein, it is easy to check that our paper satisfies the three main assumptions (Difficulty, Diversity, and Uniqueness) made by Hong and Page (2004). For instance, consider Difficulty. We use $\Phi = \{\varphi_i: i \in T\}$ or $\mathscr{P} = \{\Pi_i: i \in T\}$ to denote the team of agents in either perspective. Hong and Page (2004) state that, for any φ in Φ, there exists a solution x such that $\varphi(x) \neq x^*$. That is, for each agent, there exists at least one starting point from which the global optimum x^* is not available; any problem solver has a nut he cannot crack. Hence, the problem is difficult because no agent alone is sure to be always successful. Our Assumption 1 that no agent is endowed with the trivial partition is logically equivalent.

5. Teaming Up Does Not Ensure Success

The power of diversity demonstrated in §3 depends crucially on our assumptions about the model that describes how the problem-solving abilities of the members forming a team are drawn. Clearly, one cannot claim that Theorem 1 holds in any situation. This section demonstrates a different model under which the opposite conclusion is reached: the probability that a team (regardless of its size) is always successful at solving a problem is zero. At the end of this section, we compare the two models and gain insight in what drivers enable the formation of successful teams.

The uniform model studied in §3 assumes that agents are independently chosen according to the uniform distribution over all partitions. This section works with what we call the *urn model* after the following intuitive description. (A formal definition is given immediately after.) There are n numbered balls and n urns. One by one, each ball is tossed into a urn independently chosen with equal probability. At the end of the process, some balls end up in the same urn and some urns remain empty. Each ball represents a solution in X_n and each nonempty urn a block in the agent's partition. The procedure is repeated afresh for each agent.

For instance, suppose $n = 3$ so that there are three balls to be distributed in three urns. This can be done in $3^3 = 27$ possible ways. When the three balls end up in the same urn, we generate the trivial partition. Because there are three urns, the trivial partition is generated three times and, under the assumption that a ball is equally likely to end up in any urn, the probability of the trivial partition is $3/27 = 1/9$. Similarly, there are six possible ways to generate each of the other four possible partitions 1|23, 2|13, 3|12, and 1|2|3. Hence, the urn model assigns probability $6/27 = 2/9$ to each of these partitions.

Table 5   Probability of the Trivial Partition Under the Urn Model

	No difficulty						Difficulty				
	123	1|23	2|31	3|12	1|2|3			1|23	2|31	3|12	1|2|3
123	1/81	2/81	2/81	2/81	2/81	1/9					
1|23	2/81		4/81	4/81		2/9	1|23		1/16	1/16	1/4
2|31	2/81	4/81		4/81		2/9	2|31	1/16		1/16	1/4
3|12	2/81	4/81	4/81			2/9	3|12	1/16	1/16		1/4
1|2|3	2/81					2/9	1|2|3				1/4
	1/9	2/9	2/9	2/9	2/9			1/4	1/4	1/4	1/4	1/4

The formal definition is as follows. For a given mapping $f: X_n \to X_n$, the sets $\{x \in X: f(x) = y\}$ form a partition Π_f of X_n. If f is chosen uniformly at random from the set of all n^n mappings then Π_f is random, but not uniform. For instance, consider the trivial partition. Under the uniform model without difficulty, the probability to draw this partition over a set X_n is $1/B_n$; under the urn model without difficulty, it is $1/n^{n-1}$. Using (1), it is easy to check that $B_n < n^{n-1}$ for any $n \geq 2$, so the trivial partition is more likely to occur under the uniform model.

Similarly to §3, we build up intuition for the urn model by looking first at the case where n is quite small. Consider two agents and three elements, so that $m = 2$ and $n = 3$ with $X = \{1, 2, 3\}$. There are $B_3 = 5$ possible partitions. As discussed above, the urn model generates the trivial partition with probability $1/9$, and each of the other four with probability $2/9$.

To compute the probability that a team of two agents always discovers the global maximum, we apply these marginal probabilities to Table 1 and obtain Table 5. Its two panels are arranged as in Table 1: on the left, the case without difficulty; on the right, the case with difficulty. Adding up the joint probabilities, we find that the probability that a team of $m = 2$ agents is always successful is $41/81 \approx 0.506$ without difficulty and $3/8 = 0.375$ with difficulty.

Table 6 reports the probability that the team of $m = 2$ agents is always successful for $n = 2, 3, 4, 5$ under the urn model with (and without) difficulty. Contrary to the uniform model with difficulty, this probability is now decreasing in n—exception made for the special case at $n = 2$. When the size of the solution space grows, a team seems to become less and less effective in finding the optimal solution. Theorem 2 confirms this conjecture by proving that this probability actually goes to zero when $n \uparrow \infty$.

THEOREM 2. *Under the urn model, the probability that a team of at least $m \geq 2$ people is always successful tends to 0 as n increases.*

It is worth noting that the proof works unchanged for $m = 1$ as well, so that the conclusion of Theorem 2 actually applies for any $m \geq 1$. Moreover, similarly to Theorem 1, it holds regardless of whether or not difficulty is assumed.

Why is there such a stark difference between the uniform model and the urn model? It can be shown that, when n is large, the expected number of blocks in a random partition for the uniform model is $n/\log n$ and it is $(1 - e^{-1})n$ for the urn model (Sachkov 1997). Intuitively, the uniform model tends to generate fewer blocks than the urn model. Because the total number of solutions to partition among the available blocks is the same under both models, this implies that the average number of solutions per block is higher under the uniform model. Hence, this latter model tends to produce partitions that on average offer more bridges to a better solution. Differently put, given an arbitrary solution x, an agent with problem-solving ability drawn according to the uniform model has on average a larger set $\Pi(x)$ of solutions available; that is, he is more imaginative.

6.   Defendit Numerus

It is a reasonable assumption that the size of the solution space should adversely affect the ability of an agent to solve the problem. When n grows large, an agent should be less apt to discover the optimal solution. Our analysis for the uniform model and the urn model sidesteps an explicit discussion of this issue. This section introduces and studies a third model that formalizes this effect in a simple and attractive way.

Table 6   Probability That a Team of Two Agents Is Always Successful Under the Urn Model

	$n = 2$	$n = 3$	$n = 4$	$n = 5$
Difficulty	0	$6/16 = 0.375$	$138/441 \approx 0.313$	$5{,}400/24{,}336 \approx 0.222$
No difficulty	$3/4 = 0.750$	$41/81 = 0.506$	$1{,}369/4{,}096 \approx 0.334$	$87{,}649/390{,}625 \approx 0.224$

Roughly speaking, the model is based on the following idea. Recall that the problem-solving ability of an agent is represented by a partition Π of the solution space X. This partition describes the search space of the agent: For each $x \in X$, $\Pi(x)$ is the set of solutions that he can explore when he is given a candidate solution x. An alternative representation is by means of an undirected graph G. Let X be the set of nodes in G. Two nodes x and y in G are linked by an edge if and only if they belong to the same block in Π or, equivalently, if and only if $x \in \Pi(y)$. Then, each block of the partition Π corresponds to a connected component of the associated graph G. Intuitively, if one thinks of an edge as a mental shortcut linking a solution x to y, two nodes are in the same block if an agent contemplating x can think of y as well.

More generally, if we start with an arbitrary (undirected) graph G_n over n nodes, we can define the partition Π formed by the connected components of G_n. This second construction is more general, because it assumes that two nodes belong to the same block if there is a path connecting them (but not necessarily a direct link). Intuitively, if each edge represents a link from x to y, then x and y belong to the same block when an agent can reach and explore solution y when starting at x, through a chain of (possibly multiple) intermediate steps $z, w, \ldots$.

This insight allows the tools of graph theory to bear on the modeling of the problem-solving abilities of an agent. Formally speaking, we view a partition Π of X_n as a graph G_n with n nodes, where each block of Π is a maximally connected subgraph of G_n. Within this setting, for instance, the trivial partition Π_0 formed by the unique block X_n corresponds to the property that G_n is a connected graph.

The model studied in this section is based on the theory of random graphs and thus will be called the *random graph* model. A random graph is obtained by starting with a set of n vertices and adding edges between them at random. A celebrated approach to the generation of random graphs is due to Erdős-Rényi. Given the set X_n, each of the $\binom{n}{2}$ possible edges between distinct nodes from X_n is independently added with the same probability p. We denote the random graph by $\Gamma(n, p)$ to highlight the number n of nodes and the probability p of adding an edge.

Clearly, the structure of the partition Π associated with the random graph $\Gamma(n, p)$ depends on p: The larger p, the lower the expected number of blocks in Π. For instance, for $p = 0$ all nodes are isolated and we get the finest partition, and for $p = 1$ the graph is fully connected and we obtain the trivial partition Π_0. Thus, higher values of p correspond to higher problem-solving ability. Consistently with the objectives of this section, we assume that the size of the solution space X_n adversely affects the problem-solving abilities of an agent and therefore that p is a decreasing function of n; when useful, we write p_n to remark the dependence on n.

The random graph model postulates that the problem-solving abilities of each member of the team are represented by the partitions generated by i.i.d. draws of $\Gamma(n, p_n)$. For simplicity, we assume the same function p_n for each agent. Diversity still matters, because each agent is associated with a different random graph. Therefore, two nodes x and y may be linked by an edge (and be available from each other) for one agent but not for another. This leads to the key observation that uncovers the power of a diverse team.

Given two nodes x and y, the probability that a single agent connects them is p. However, it is enough that *one* of the agents has an edge between two nodes to make them available to the entire team. The probability that at least one of m teammates connects x and y is $1 - (1 - p)^m$. Thus, we can build the random graph representing the team's problem-solving ability by assuming that each edge between two nodes is independently present with probability $1 - (1 - p)^m$. Formally speaking, when the partition of each of the m agents in a team is represented by a random graph $\Gamma(n, p)$, the meet is described by the random graph $\Gamma(n, 1 - (1 - p)^m)$. When this graph is connected, the meet is the trivial partition and the team has access to the whole solution space, making sure that the solution is always found.

As usual, our main object of interest is the probability that a team is always successful when n is large. Therefore, we need to consider the probability that $\Gamma(n, 1 - (1 - p)^m)$ is connected. An important result from random graph theory provides a powerful sufficient condition. If (eventually) $p_n \geq (1 + \varepsilon) \ln n / n$ for some $\varepsilon > 0$, the probability that $\Gamma(n, p_n)$ is connected tends to 1 as $n \uparrow \infty$; and vice versa, if (eventually) $p_n \leq (1 - \varepsilon) \ln n / n$, then this probability tends to 0; see Theorem 2.8.1 in Durrett (2007). This result states that the connectedness of the graph undergoes a phase transition around the sharp threshold provided by $\ln n / n$. Given a team of m agents whose meet is described by the random graph $\Gamma(n, 1 - (1 - p_n)^m)$, this implies the following result.

Theorem 3. *Under the random graph model, the probability that a team of at least $m \geq 1$ people is always successful tends to 1 as n increases if*

$$1 - (1 - p_n)^m \geq (1 + \varepsilon) \frac{\ln n}{n} \tag{2}$$

for all sufficiently large n, and tends to 0 if

$$1 - (1 - p_n)^m \leq (1 - \varepsilon) \frac{\ln n}{n} \tag{3}$$

for all sufficiently large n.

Table 7	Minimum Size m for an (Almost) Always Successful Team as a Function of n			
n	p_n	$\ln n/n$	$1 - (1 - p_n)^m$	m
10	0.111111	0.230259	0.297668	3
10^2	0.010101	0.046052	0.046052	5
10^3	0.001001	0.006908	0.006986	7
10^4	0.000100	0.000921	0.001000	10
10^5	0.000010	0.000115	0.000120	12
10^6	0.000001	0.000014	0.000014	14

This sufficient condition extends easily to the case where we remove the simplifying assumption that the random graph for each agent in the team is drawn independently according to the same model $\Gamma(n, p_n)$. In fact, suppose that the partition for each agent i is drawn according to a different model $\Gamma(n, p_n^i)$. Define $\hat{p}_n = \min_i p_n^i$ and substitute it for p_n in (2). Similarly, it would suffice to replace $\bar{p}_n = \max_i p_n^i$ in (3).

Theorem 3 formalizes an important trade-off, that we illustrate by means of an example. Suppose that $p_n = 1/(n-1)$ for each agent in the team. Fix n and an arbitrary node x. Because there are other $n-1$ nodes besides it, node x can have at most $n-1$ incident edges. Each of these edges has probability p_n of being present, so the expected number of edges incident to x (a.k.a. *degree* of x) follows a binomial distribution $B(n-1, p_n)$. Hence, $p_n = 1/(n-1)$ corresponds to the special case where each node has an expected degree of $(n-1)p_n = 1$ or each agent has access on average to only another solution from any node x. This makes his availability map pretty limited and, indeed, applying (3) shows the probability that a single agent is always successful, for large n must be close to zero.

However, if we put together a sufficiently large number m of agents, we can make the probability that they will be jointly successful very close to one as far as (2) holds. Table 7 shows the minimum size of a team that is necessary to pass the threshold $\ln n/n$ and make the team very likely to be always successful for $n = 10, 10^2, 10^3, 10^4, 10^5, 10^6$ and $p_n = 1/(n-1)$. All values are approximated to the sixth decimal digit.

As it is easy to see, the minimum number of members necessary to make the team very likely to be always successful increases very slowly with respect to the size of the solution space. This is not a coincidence.[1] Because p_n is very small, we can use the Taylor's approximation $1 - (1 - p_n)^m \approx mp_n$ and rewrite (2) as $mp_n = \ln n/n$, so that $m \approx [(n-1)/n] \ln n \approx \ln n$. In other words, the minimum size for an always successful team increases logarithmically with respect to the size of the solution space. The power of diversity is apparent. Each agent in the team has an

expected degree of one so his problem-solving ability is severely limited. However, even a small team can span a huge search space because the number of possible combinations across the perspectives of its agents increases exponentially. There is power in (even small) numbers or, for short, *defendit numerus*.

The intuition that a small increase in a team makes it more effective at finding the solution is also true for the uniform and for the urn model. However, because these models exhibit no phase transition, the limit of the probability that a team is always successful as $n \uparrow \infty$ is the same, regardless of the number of its members. That is, the size of the team affects the speed of convergence but not the value of the limit. The phase transition exhibited by the random graph model, instead, implies that the size of the team may radically change the limit value from zero to one.

Nonetheless, Collevecchio and LiCalzi (2011) prove that, if the number of urns may be different from the number of balls, the urn model can be nested in a more general scheme that exhibits a phase transition. Reinterpreted in our framework, their result implies that the probability that a team is always successful goes to zero or one depending both on the number and on the cognitive ability of the agents.

7. Groupthink Impairs the Power of Diversity

Groupthink is a situation where the members of a team align their thinking. Originally coined by Whyte (1952), the term entered the academic literature in Janis (1972), where it was argued that the search for consensus in the decision-making team had adversely affected the quality of some important policy decisions. In the psychological literature, groupthink carries a pejorative connotation and it is sometimes offered as a possible explanation for why individuals might make better decisions than groups. Within this paper, we take a neutral stance over the causes and effects of groupthink except for its impact on the ability of a team to find the optimal solution.

To study how groupthink affects team performance, we enrich the random graph model introduced in the previous section with the assumption that the probability p_n that an agent has an edge between two nodes x and y is positively dependent among agents. Intuitively, the higher the groupthink, the higher the correlation. Consider first the two extreme cases. Under independence, the probability that a team of m agents has an edge between two nodes is $1 - (1 - p_n)^m$, as discussed in the previous section. Under complete dependence, all agents think exactly alike so the probability that an edge is present is p_n, independently of the number of agents in the team.

To study the intermediate cases, we model the dependence across agents as a convex combination

---

[1] Laughlin et al. (2006) provide empirical evidence that the marginal contribution from the size of the team is decreasing.

of the two extreme cases. Suppose that each agent has the same marginal probability p_n to have an edge between any two nodes. Then the joint probability that the edge is present for a team of m agents is

$$\delta p_n + (1-\delta)[1-(1-p_n)^m], \tag{4}$$

where δ in $[0,1]$ is a dependence parameter that describes the alignment between agents' partitions. We call this the *random graph model with dependence*. For $\delta = 0$, we obtain the random graph model with independence among agents; for $\delta = 1$, we obtain the random graph model with complete dependence. Clearly, there are many alternative ways to introduce dependence in the random graph model; however, our choice is the simplest one that suffices to capture the features in which we are interested. See the B11 copula in Joe (1997) for its properties. The next result follows immediately from (4) and the techniques of §6.

THEOREM 4. *Under the random graph model with dependence, the probability that a team of at least $m \geq 1$ people is always successful tends to 1 as n increases if*

$$\delta p_n + (1-\delta)[1-(1-p_n)^m] \geq (1+\varepsilon)\frac{\ln n}{n} \tag{5}$$

for all sufficiently large n, and tends to 0 if

$$\delta p_n + (1-\delta)[1-(1-p_n)^m] \leq (1-\varepsilon)\frac{\ln n}{n} \tag{6}$$

for all sufficiently large n.

Theorem 4 can be used to compare how more groupthink (represented by a higher value of δ) affects the performance of a team over a large solution space. The following example illustrates how. As in §6, suppose that $p_n = 1/(n-1)$ so that each node has an expected degree of one. Under independence, for $\delta = 0$, we have seen that a single agent may fail to find the optimal solution; however, we can assemble a sufficiently large team to ensure success. Table 8 reports the minimum size m of a team necessary to satisfy (5) for a given value of n (and a sufficiently small $\varepsilon > 0$) to make the team almost sure to be always successful for $\delta = 0, 0.25, 0.50, 0.75, 1$ and

$n = 10, 10^2, 10^3, 10^4, 10^5, 10^6$. The first four columns (up to $\delta = 0$) are the same as in Table 7.

The focus of interest is the relationship between m and δ. It is easy to see that m is increasing in δ: The greater the extent of groupthink, the larger the number of agents for a successful team. When groupthink is extreme and $\delta = 1$, any team is as good as a single agent, so no size is sufficiently large to ensure success. More interestingly, if we treat m as a continuous variable for the sake of convenience, it is easy to see that the relationship is also convex. When δ increases from 0 to 0.25, a modest upward adjustment in m suffices to preserve the performance of the team. However, when δ goes from 0.50 to 0.75, a much larger adjustment is necessary.

An approximate but effective estimate of the phenomenon can be obtained as follows. Use the Taylor's approximation $1 - (1-p_n)^m \approx mp_n$ and substitute in (5) to obtain $\delta p_n + (1-\delta)mp_n > \ln n/n$, or

$$m > \frac{\ln n/n - \delta p_n}{(1-\delta)p_n}.$$

For n large, $p_n \cdot n \approx 1$ and $\ln n \gg \delta$ so the right-hand side approximates to $\ln n/(1-\delta)$. Under no groupthink, when $\delta = 0$, we know already that the minimum size for a successful team increases logarithmically in the dimension of the solution space. However, under groupthink, this logarithmic growth is magnified by the $1/(1-\delta)$ multiplicative factor. As $\delta \uparrow 1$, the necessary size increases without bounds. This shows clearly the countervailing effect of groupthink on diversity: the more likely the agents are to have the same perspective on a problem, the greater is the number of agents required to cooperate before the optimal solution can be found.

Although it is plausible to expect that correlations exist among the perspectives used by teammates, it is important that they should not be so high to wipe out the advantages brought by the diversity in perspectives (Lazear 1999). Clearly, this intuition carries over to other models, as long as we view groupthink as a form of positive dependence among agents' partitions. Nonetheless, it is an open problem to find a simple and tractable way to formalize positive dependence in the uniform or in the urn models.

8. Closing Comments

This paper studies the power of diversity in a model where a team of agents with limited problem-solving ability faces a disjunctive task over a large solution space. From a managerial viewpoint, the overarching conclusion is that firms searching solutions over difficult intellective problems need to tap people with different toolboxes.

From a modeling perspective, it is important to assess the reach of our model. We formalize the limitations in agents' abilities using partitions and the

Table 8   Minimum Size m for an (Almost) Always Successful Team Under Groupthink

			m				
n	p_n	$\ln n/n$	$\delta = 0$	$\delta = 0.25$	$\delta = 0.50$	$\delta = 0.75$	$\delta = 1$
10	0.111111	0.230259	3	3	4	8	∞
10^2	0.010101	0.046052	5	6	9	17	∞
10^3	0.001001	0.006908	7	9	13	25	∞
10^4	0.000100	0.000921	10	12	18	34	∞
10^5	0.000010	0.000115	12	15	22	43	∞
10^6	0.000001	0.000014	14	18	26	51	∞

variability in their perspectives as randomness over the partitions. As discussed in §4, this is formally equivalent to a setup (Hong and Page 2001) where agents draw heuristics from their toolbox. However, the assumption that agents' partitions are independently drawn from the same distribution is a crude approximation and cannot capture the degree with which agents' toolboxes overlap in different organizations, making them more or less apt to problem solving. It may be appropriate to qualify this special case as a form of "idiosyncratic" search, where each agent's toolbox is independent from his teammates'.

A useful benchmark is the compact summary in Lazear (1999) of three main factors driving the success of a diverse team working on an intellective task: (a) how disjoint are the agents' information sets; (b) how relevant they are; and (c) the quality of the communication processes. Our model implicitly recognizes and takes into account the first two items, but assumes away any negative effect from the third one. Communication is advantageous because it may unstuck a team from a bad solution. However, many different biases may adversely affect knowledge transfer among agents (Argote et al. 2000). Diversity promotes both improved knowledge sharing (Cummings 2004) and more misunderstandings (Postrel 2002).

A different line of research with important analogies to our model is the study of groups engaged in idea generation, where performance is defined as the quality of the best ideas identified (Girotra et al. 2010). These processes can exhibit several inefficiencies (Paulus and Yang 2000). In particular, when search is carried out in parallel, they are subject to redundancies that may become very expensive over large spaces (Kornish and Ulrich 2011). Our model will benefit when adequate formalizations for the effects of communication and for the process of idea sharing shall be available.

Acknowledgments

Appendix

PROOF OF THEOREM 1. As discussed above, the probability that a team is always successful is equal to the probability that the meet of their partitions is the trivial partition. Let P_{mn} denote the probability that the meet of $m \geq 2$ partitions of X_n chosen under the uniform model with difficulty is the trivial partition. It suffices to show that $P_{mn} \to 1$ as $n \uparrow \infty$.

Denote by P'_{mn} the probability that the meet is the trivial partition under the uniform model without difficulty. Theorem 5 in Pittel (2000) shows that

$$P'_{mn} = 1 - O\left(\frac{\log^{m+1} n}{n^{m-1}}\right) \tag{7}$$

from which it follows immediately that $\lim_{n \uparrow \infty} P'_{mn} = 1$. We establish the claim by proving that $(1 - P'_{mn})/(1 - P_{mn}) \to 1$ as $n \uparrow \infty$.

The difficulty assumption removes the trivial partition from the set of possible partitions of each agent. This decreases by one the number of equally likely partitions for each agent, from B_n to $B_n - 1$. On the other hand, this removal does not affect the total number of m-tuples of partitions whose meet is *not* the trivial partition. Denote this number by W_n. Then

$$\lim_{n \uparrow \infty} \frac{1 - P'_{mn}}{1 - P_{mn}} = \frac{W_n/(B_n)^m}{W_n/(B_n - 1)^m} = \left(\frac{B_n - 1}{B_n}\right)^m = 1.$$

Note also that $P'_{mn} \to 1$ by Pittel's (2000) result, so Theorem 1 holds for the uniform model regardless of whether or not difficulty is assumed. □

PROOF OF THEOREM 2. Let P'_{mn} denote the probability that the meet is the trivial partition under the urn model without difficulty. Because $P'_{mn} \geq P_{mn}$, it suffices to show that $P'_{mn} \to 0$ and the theorem holds regardless of whether or not we assume difficulty. The strategy of the proof is the following.

We say that a solution j is *isolated* for an agent i when $\{j\}$ is a singleton block for his partition. Analogously, a solution j is isolated for the team if it is isolated for each agent. Let A^i_j and A_j denote the event that the jth solution is isolated for agent i and for the team, respectively. When the meet of the agents' partitions is the trivial partition, no solution j can be isolated for the team. Therefore,

$$P'_{mn} \leq 1 - P\left(\bigcup_{j=1}^n A_j\right).$$

We are going to show that $P(\bigcup_{j=1}^n A_j) \to 1$ as $n \uparrow \infty$.

We begin with a few preliminary observations. Given a set X_n of n possible solutions, A^i_j corresponds to the event that the jth ball ends up alone in one of the n urns; thus,

$$P(A^i_j) = \frac{n(n-1)^{n-1}}{n^n} = \left(1 - \frac{1}{n}\right)^{n-1}.$$

Analogously, for $j_1 < j_2 \cdots < j_k$, the probability that $k \leq n$ solutions in X_n are isolated for agent i is

$$P\left(\bigcap_{s=1}^k A^i_{j_s}\right) = \prod_{s=1}^k \left(1 - \frac{1}{n-s+1}\right)^{n-s}.$$

To see why, use the following inductive argument. The probability that the j_1th solution is isolated (or, equivalently, the j_1th ball ends up alone) is $(1 - 1/n)^{(n-1)}$. Conditional on this event, the other $n - 1$ balls are distributed uniformly in the remaining $n - 1$ urns. Hence, the probability that the j_2th solution is isolated (or, equivalently, the j_2th ball ends up alone) is $(1 - 1/(n - 1))^{(n-2)}$. And so on.

Finally, because all agents' partitions are identically and independently distributed,

$$P(A_j) = \prod_{i=1}^{m} P(A_j^i) = (P(A_j^i))^m$$

and thus, for $j_1 < j_2 \cdots < j_k$, the probability that $k \le n$ solutions in X_n are isolated for the team is

$$P\left(\bigcap_{s=1}^{k} A_{j_s}\right) = P\left(\bigcap_{i=1}^{m}\bigcap_{s=1}^{k} A_{j_s}^i\right) = \prod_{s=1}^{k}\left(1 - \frac{1}{n-s+1}\right)^{m(n-s)}. \quad (8)$$

Note that, for any fixed k, $P(\bigcap_{s=1}^{k} A_{j_s}) \to e^{-mk}$ as $n \uparrow \infty$.

We are now ready for the main argument. Given $\varepsilon > 0$, choose a sufficiently large integer M so that $(1 - e^{-m})^M < \varepsilon$. By the inclusion–exclusion formula, we have

$$P\left(\bigcup_{j=1}^{M} A_j\right) = \left[\sum_{j=1}^{M} P(A_j) - \sum_{1 \le j_1 < j_2 \le M} P(A_{j_1} \cap A_{j_2}) + \cdots \right.$$
$$\left. + (-1)^{M+1} P\left(\bigcap_{j=1}^{M} A_j\right)\right].$$

Taking limits on both sides,

$$\lim_n P\left(\bigcup_{j=1}^{M} A_j\right)$$
$$= \sum_{j=1}^{M}(-1)^{j+1}\binom{M}{j}e^{-mj} = -\sum_{j=1}^{M}\binom{M}{j}(-e^{-m})^j$$
$$= 1 - \sum_{j=0}^{M}\binom{M}{j}(-e^{-m})^j = 1 - \sum_{j=0}^{M}\binom{M}{j}(-e^{-m})^j \cdot 1^{M-j}$$
$$= 1 - (1 - e^{-m})^M > 1 - \varepsilon,$$

where the last step follows from our choice of M. This concludes the proof. □

References

Argote, L., P. Ingram, J. M. Levine, R. L. Moreland. 2000. Knowledge transfer in organizations: Learning from the experience of others. *Organ. Behav. Human Decisions* **82** 1–8.

Berend, D., T. Tassa. 2010. Improved bounds on Bell numbers and on moments of sums of random variables. *Probab. Math. Statist.* **30** 185–205.

Blinder, A. S., J. Morgan. 2005. Are two heads better than one? Monetary policy by committee. *J. Money, Credit, Banking* **37** 798–811.

Blinder, A. S., J. Morgan. 2008. Do monetary policy committees need leaders? A report on an experiment. *Amer. Econom. Rev.* **98** 224–229.

Bornstein, G., T. Kugler, A. Ziegelmeyer. 2004. Individual and group decisions in the centipede game: Are groups more "rational" players? *J. Exp. Soc. Psych.* **40** 599–605.

Burns, J. G., A. G. Dyer. 2004. Diversity of speed-accuracy strategies benefits social insects. *Current Biology* **18** R953–R954.

Collevecchio, A., M. LiCalzi. 2011. The probability of nontrivial common knowledge. Working Paper 6/2011, Università Ca' Foscari Venezia, Venezia, Italy.

Conradt, L., C. List. 2009. Group decisions in humans and animals: A survey. *Philos. Trans. Royal Soc. B* **364** 719–742.

Cooper, D. J., J. H. Kagel. 2005. Are two heads better than one? Team versus individual play in signaling games. *Amer. Econom. Rev.* **95** 477–509.

Cummings, J. N. 2004. Work groups, structural diversity, and knowledge sharing in a global organization. *Management Sci.* **50** 352–364.

Dahlin, K. B., L. R. Weingart, P. J. Hinds. 2005. Team diversity and information use. *Acad. Management J.* **48** 1107–1123.

Davis, J. H. 1992. Some compelling intuitions about group consensus decisions, theoretical and empirical research, and interpersonal aggregation phenomena: Selected examples, 1950–1990. *Organ. Behav. Human Decisions* **52** 3–38.

Dowling, D. W. 2007. The best advice I ever got: Hans-Paul Bürkner, president and chief executive officer, the Boston Consulting Group. *Harvard Bus. Rev.* (December), http://hbr.org/2007/12/the-best-advice-i-ever-got-hans-paul-burkner-president-and-chief-executive-officer-the-boston-consulting-group/ar/1.

Durrett, R. 2007. *Random Graph Dynamics*. Cambridge University Press, Cambridge, UK.

Feynman, R., R. Leighton. 1985. *Surely You're Joking, Mr. Feynman!: Adventures of a Curious Character*. W.W. Norton, New York.

Gerchak, Y., Q.-M. He. 2004. Group size and parallelism effects in tasks with heterogeneous levels of difficulty: A stochastic order approach. *J. Systems Sci. Systems Engrg.* **13** 257–278.

Girotra, K., C. Terwiesch, K. T. Ulrich. 2010. Idea generation and the quality of the best idea. *Management Sci.* **56** 591–605.

Hastie, R. 1986. Review essay: Experimental evidence on group accuracy. G. Owen, B. Grofman, eds. *Information Pooling and Group Decision Making*. JAI Press, Greenwich, CT, 129–157.

Hill, G. W. 1982. Group versus individual performance: Are $N+1$ heads better than one? *Psych. Bull.* **91** 517–539.

Hong, L., S. Page. 2001. Problem solving by heterogeneous agents. *J. Econom. Theory* **97** 123–163.

Hong, L., S. Page. 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proc. Natl. Acad. Sci. USA* **101** 16385–16389.

Jackson, S. E., A. Joshi, N. L. Erhardt. 2003. Recent research on team and organizational diversity: SWOT analysis and implications. *J. Management* **29** 801–830.

Janis, I. 1972. *Victims of Groupthink: Psychological Studies of Policy Decisions and Fiascoes*. Houghton Mifflin, Boston.

Joe, H. 1997. *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.

Kocher, M. G., M. Sutter. 2005. The decision maker matters: Individual versus group behavior in experimental beauty-contest games. *Econom. J.* **115** 200–223.

Kornik, J. 2008. Excellence in diversity: Consulting firms honored. *Consulting Magazine* (July 29), http://www.consultingmag.com/article/ARTALPII792?C=CeHrcuiZ2tPzIC90.

Kornish, K. J., K. T. Ulrich. 2011. Opportunity spaces in innovation: Empirical analysis of large samples of ideas. *Management Sci.* **57** 107–128.

Krishnan, H. A., A. Miller, W. Q. Judge. 1997. Diversification and top management team complementarity: Is performance improved by merging similar or dissimilar teams? *Strategic Management J.* **18** 361–374.

Laughlin, P. R. 1999. Collective induction: Twelve postulates. *Organ. Behav. Human Decisions* **80** 50–69.

Laughlin, P. R., B. L. Bonner, A. G. Miner. 2002. Groups perform better than the best individuals on letters-to-numbers problems. *Organ. Behav. Human Decisions* **88** 605–620.

Laughlin, P. R., E. C. Hatch, J. S. Silver, L. Boh. 2006. Groups perform better than the best individuals on letters-to-numbers problems: Effects of group size. *J. Personality Soc. Psych.* **90** 644–651.

Lazear, E. P. 1999. Globalisation and the market for teammates. *Econom. J.* **109** C15–C40.

Levine, J. M., R. L. Moreland. 1998. Small groups. D. T. Gilbert, S. T. Fiske, G. Lindzey, eds. *Handbook of Social Psychology*, Vol. 2. McGraw-Hill, Boston, 415–469.

Liker, A., V. Bókony. 2009. Larger groups are more successful in innovative problem solving in house sparrows. *Proc. Natl. Acad. Sci. USA* **106** 7893–7898.

Mannix, E., M. A. Neale. 2005. What differences make a difference? The promise and reality of diverse teams in organizations. *Psych. Sci. Public Interest* **6** 31–55.

Marschak, J., R. Radner. 1972. *Economic Theory of Teams*. Yale University Press, New Haven, CT.

Michaelsen, L. K., W. E. Watson, R. H. Black. 1989. A realistic test of individual versus group consensus decision making. *J. Appl. Psych.* **74** 834–839.

Page, S. 2007. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, Societies*. Princeton University Press, Princeton, NJ.

Paulus, P. B., H. -C. Yang. 2000. Idea generation in groups: A basis for creativity in organizations. *Organ. Behav. Human Decisions* **82** 76–87.

Pelled, L. H., K. M. Eisenhardt, K. R. Xin. 1999. Exploring the black box: An analysis of work group diversity, conflict, and performance. *Admin. Sci. Quart.* **44** 1–28.

Pitman, J. 2006. *Combinatorial Stochastic Processes*. Springer-Verlag, Berlin.

Pittel, B. 2000. Where the typical set partitions meet and join. *Electronic J. Combinatorics* **7** R5.

Postrel, S. 2002. Islands of shared knowledge: Specialization and mutual understanding in problem-solving teams. *Organ. Sci.* **13** 303–320.

Reagans, R., E. Zuckerman, B. McEvily. 2004. How to make the team: Social networks vs. demography as criteria for designing effective teams. *Admin. Sci. Quart.* **49** 101–133.

Rubinstein, A. 1993. On price recognition and computational complexity in a monopolistic model. *J. Political Econom.* **101** 473–484.

Sachkov, V. N. 1997. *Probabilistic Methods in Combinatorial Analysis*. Cambridge University Press, Cambridge, UK.

Simons, T., L. H. Pelled, K. A. Smith. 1999. Making use of difference: Diversity, debate, and decision comprehensiveness in top management teams. *Acad. Management J.* **42** 662–673.

Ward, A. J. W., J. E. Herbert-Read, D. J. T. Sumpter, J. Krause. 2011. Fast and accurate decisions through collective vigilance in fish shoals. *Proc. Natl. Acad. Sci. USA* **108** 2312–2315.

Whyte, W. H., Jr. 1952. Groupthink. *Fortune* (March) 114–117, 142, 146.