Xiaofan Liang
Eric Bonabeau
Minerva Schools at KGI
Disclaimer: this is a report for a research project, but is still incompleted

ABSTRACT

A challenge that a human resource manager might run into is to balance between the diversity of the new recruits and their abilities. Many argued that they are not opposed to diversity, but they are just not willing to sacrifice team performance. To test this statement, unravel the dynamics between diversity and ability and find an optimized balance between the two, we developed a model with the software Netlogo to simulate the impact of different weights of diversity and ability in recruiting strategies on team performance under different task sets. We found that teams perform well in tasks that correspond to their team types, but the best tradeoff balance between diversity and ability is under the group portfolio type of team-forming strategy, which created the most efficient and robust team across a variety of task sets. The time distributions of team performance reveal that adding diverse members into the team can drastically stabilize team performance across a variety to task, while solely optimizing for the "smartness" of the team members may lead to rare but significant underperformance. We also found that adjusting the sufficient level of solution and recruiting extra person solely for the purpose of diversity does not impact the results we got.

INTRODUCTION

Diversity was usually discussed as an ethnic issue for fairness, but a growing trend in social science, complex system and business literature had started to address its benefits for the collective. The definition of diversity shifted from a socio-biological focus such as gender, ethnicity and cultural background to functional focus such as team role and problem-solving heuristics to better explain its correlation with group performance. Nonetheless, group performance itself was not able to be attributed to a single factor or a clear mechanism of the interactions between different factors. Woolley's research (2010) further proved the team performance is an emergent property that is independent of individual intelligence of the group members. This implies that the traditional "picking the smartest" type of talent recruitment strategy in companies might fail since it assumed a linear sum up of individual talents to maximize group performance. Another line of research that converged on the same conclusion is Belbin's study on Apollo Syndrome (1981), which described his observation on unexpected poor group performance results from the sharp minded and analytical team. Hong and Page (2004), developed a mathematical model to further confirmed this observation that a random team, which is assumed to have better functional diversity, can systematically outperformed a high ability team. Their research also first elaborated a tension between ability and diversity that seems to mutually exclusive since when the pool of problem solvers grows larger, the best become like-minded. However, critics had questioned not only Hong and Page's mathematical assumptions in the model but also the real life implications of their results. Hong and Page broke the promise of ability but did not provide any specifications on the extent diversity trumps ability and what will be a more realistic solution for human resource managers other than random recruitment (Thompson, 2014; Josh, 2014).

This paper builds on the critics for Hong and Page's research to retest the team performance of different team types (not limited to random and high ability team) with a Netlogo modeling framework. In addition, we contextualize team performance with different task sets and explore an optimal balance between diversity and technical ability. We believe that team performance is the interaction of three variables - technical and cognitive abilities (refers to as tech ability), communication costs, and perspectives. Technical and cognitive abilities are fundamental and necessary to solve analytical and skill-based problems. Communication costs are time people spend talking to each other in the project, which we assume it to be a constraint to diverse team since it will have more social frictions due to different background. Perspectives, which is the variable we use to represent diversity, represent different ways that people think about and approach the problem. We will first go through the mechanism and specification of Netlogo model, which will provide legitimacy for the data it generated. Then we present the result in graphs and discuss the interpretations.

MODEL AND METHOD

Our Netlogo model includes two type of agents. One of them is human, but each with different index scores ranging from 1 to 5 among three key attributes that we think will impact team performance. Index scores for perspectives represent different kinds of perspectives, while those for others are interval variables representing ability from low to high. The other one is task, which composes of a collection of two things - the level of technical abilities recommended and a list of perspectives needed to solve the task. Parameters we control to generate data include team types, which determines the distribution of people drawn to form a team, and task type, which determines the distributions of components for the tasks. The mechanism between teams and tasks goes like the follow:

1) When the simulation starts, computer will draw five individuals X with distribution of technical ability, T, communication skills, C, and perspectives, P, determined by team types selected.

$$X_i = \{T_i, C_i, P_i\}, \ \ T_i, C_i, P_i \in [1,5]$$

2) After the draw, the individuals form teams that have team attributes set Z that will be used to calculate the time to finish certain tasks.

$$Z_j = \{\tfrac{1}{5}\sum_{i=1}^{5} T_i, \ \tfrac{1}{5}\sum_{i=1}^{5} C_i, \ \{P_1, P_2, P_3, P_4, P_5\}\}$$

3) Tasks selected will have attributes set K given by the task type that will also be used to calculate the time to finish certain tasks.

$$K_i = \{T_i, C_i, \{P_1, P_2, \ldots P_m\}\}, \ m \in [1,5]$$

4) The time t' needed to finish task K_i given team Z_j is the sum of time needed for each of the three attributes of Z_j to match the requirement from K_i. To make sure that three factors have equal impacts on the time performance, we devised two different ways to calculate those that are interval variables (tech ability & communication) and perspective which is categorical, so that the distribution of time for each factor is roughly the same.

$$t' = t_T' + t_C' + t_P'$$

4.1) Time to match perspectives, t_P is calculated through a simulation of problem solving as searching through a circle with seven dots. All team member start at the same dot, while different perspectives, denoted from one to five, representing different steps that each person will take each tick to go around the circle. Solution can require between one and five perspectives, each also denoted as number from one to five, will point to the location away from the starting point as denoted number. Prime number seven is chosen as the size of circle to ideally equalize the probability for each perspective to be advantageous given certain solution perspective. All team members will run the circle at the same time, and as long as one person has found the solution, the running will stop and count the ticks.

$$t_P' = Min\ (t_{P_1}, t_{P_2}, t_{P_3}, t_{P_4}, t_{P_5})$$

4.2) To ensure the time distribution of t_T is similar to t_P , we calculate t_T given its distance from the average level tech skills (which is 3) and scaled by a normal distribution of time derived from t_P . The mean (25) and std (6.35) of the normal distribution are taken from the distribution of perspective time with 1000 iterations of problem solving.

$$t_T' = 25\ -\ (t_T - 3)\ *\ 6.35$$

4.3) Similar to t_T , t_C is calculated given its distance from the average level communication skills (which is 3) and scaled by the same normal distribution of time derived from t_P . The only difference is that t_C also accounts for cost for diversity, which is an extra term added only if the diversity index D of a team is greater or equal to four (calculation of D will be explained in next section). This extra time is calculated as the value given its distance from mean diversity index (which is 3) and scaled by a normal distribution of time. The mean of this normal distribution is 4.2 and the std of the distribution is 1 which is the same as one std change of technical index. Here the cost of diversity is conceptualized as the communication cost and is the same as one unit cost of tech-skill index.

$$t_T' = 25 - (t_T - 3) * 6.35 \qquad if\ D < 4$$
$$t_T' = 25 - (t_T - 3) * 6.35 + (D - 3) * x \sim N(4.2, 1) \qquad if\ D \geq 4$$

Assumptions
We do have a few assumptions build into the model. For the model to mimic the real life, we built in constraints for both diversity and ability in team performance. Therefore, the distribution of perspectives for people who have high technical and cognitive skills are much narrower, which means teams formed with high technical people are less likely to be diverse in perspectives. This assumption is also validated in literature (Hong & Page, 2004). To balance this constraint, we added a cost of diversity term to communication cost for teams that display more than average diversity. Other assumptions include simplification that companies in real life may not be able to achieve. For example, we have large enough candidate pool to ensure each team type can be constructed, or team performance can be calculated through linear combination of time contributed by each factor.

RESULT

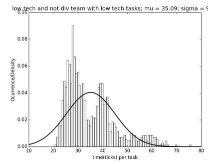
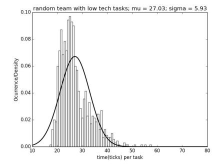
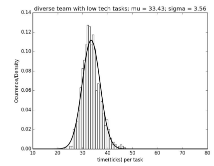
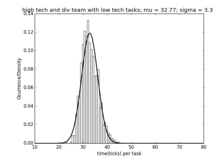
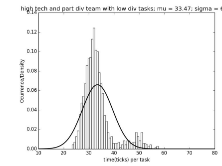
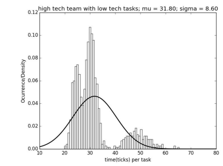
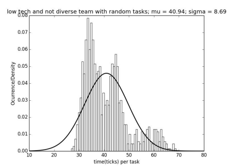
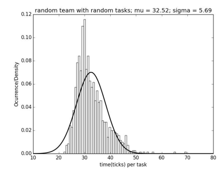
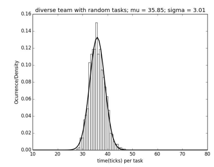
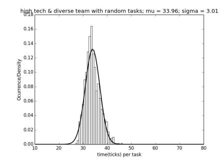
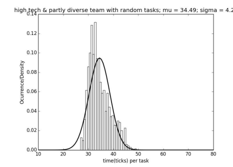
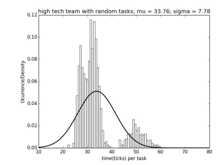
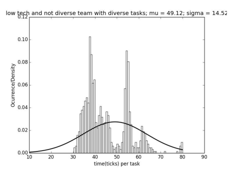
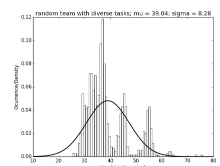
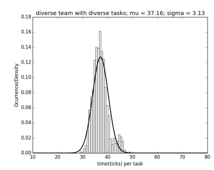
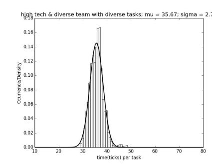
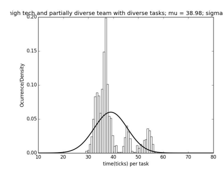
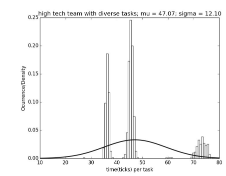
All the data being analyzed had been generated from Netlogo based on the interaction mechanism I discussed above. To make it easier to compare between team types, we created a diversity index and technical index to each team type with following definition:

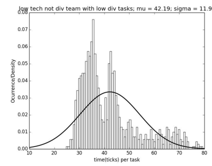
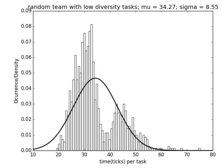
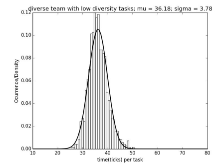
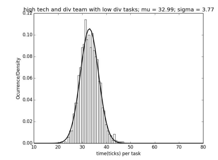
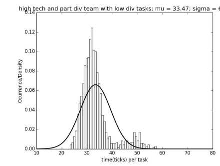
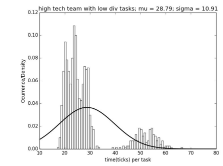
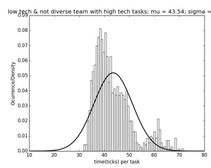
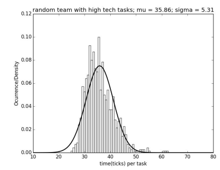
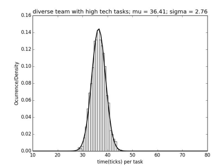
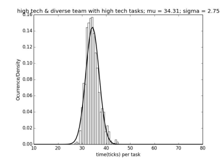
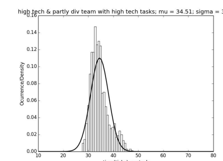
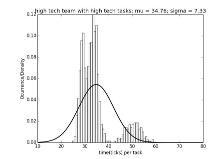
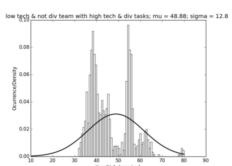
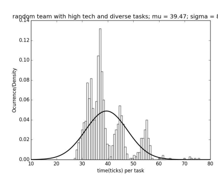
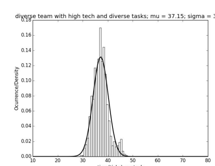
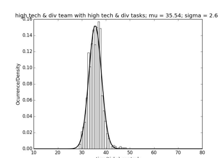
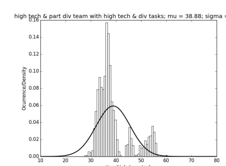
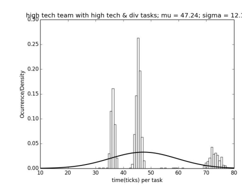
Diversity Index:
An index given to a team based on its perspectives and roles compositions, calculated as the reverse of the mean time(s) a number appears in the list of team perspectives. The more different team members' perspectives and roles are with each other, the higher the diversity index. For example, the diversity index of a team perspectives [3, 3, 4, 5, 5, 5] will be 5 - [(2 + 2 + 1 + 3 + 3 + 3) / 5].

Technical Index:
An index given to a team based on its members' average technical abilities. Its range is [1, 5].

Here is a table for the frequency distribution of team performance (measured in time units) under different team types and task sets in 1000 simulations. To compare, the x-axis of all graphs are kept in the same range of [20, 90]. For each distribution, normal distribution parameters are used as estimands for their overall performance.

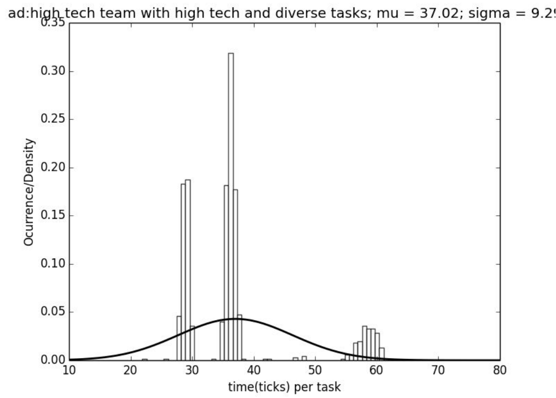
Team type	**Low-tech abilities +**	**Random**	**Diverse**	**High-tech abilities**	**High-tech abilities +**	**High-tech abilities**

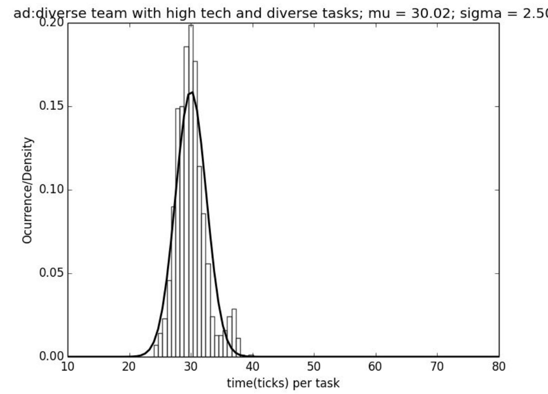
(row) **Task set** **(column)**	**not diverse** (technical index = 2.11, diversity index = 2.59)	(technical index = 3, diversity index = 3.4)	(technical index = 3.01, diversity index = 4.26)	**+ Diverse** (technical index = 3.83, diversity index = 4.18)	**Partially diverse** (technical index = 4.03, diversity index = 3.11)	(technical index = 4.49, diversity index = 2.13)
Low Tech Skills tech index = 1.5	 Mu is 35.09 Sigma is 9.87 Rank: 6	 Mu is 27.03 Sigma is 5.93 Rank: 1	 Mu is 33.43 Sigma is 3.56 Rank: 5	 Mu is 32.77 Sigma is 3.34 Rank: 3	 Mu is 33.10 Sigma is 4.19 Rank: 4	 Mu is 31.80 Sigma is 8.6 Rank: 2
Random tasks Tech index = 3	 Mu is 40.94 Sigma is 8.69 Rank: 6	 Mu is 32.52 Sigma is 5.69 Rank: 1	 Mu is 35.85 Sigma is 3.01 Rank: 5	 Mu is 33.96 Sigma is 3.01 Rank: 3	 Mu is 34.49 Sigma is 4.21 Rank: 4	 Mu is 33.76 Sigma is 7.78 Rank: 2
Diverse tasks Tech index = 3	 Mu is 49.12 Sigma is 14.52 Rank: 6	 Mu is 39.04 Sigma is 8.28 Rank: 4	 Mu is 37.16 Sigma is 3.13 Rank: 2	 Mu is 35.67 Sigma is 2.73 Rank: 1	 Mu = 38.98 Sigma = 6.64 Rank: 3	 Mu is 47.07 Sigma is 12.10 Rank: 5

Low Diversity tasks (tasks that favor tech people's perspectives) Tech index = 3	 Mu is 42.19 Sigma is 11.92 Rank: 6	 Mu is 34.27 Sigma is 8.55 Rank: 4	 Mu is 36.18 Sigma is 3.78 Rank: 5	 Mu is 32.99 Sigma is 3.77 Rank: 2	 Mu = 33.47 Sigma = 6.04 Rank: 3	 Mu is 28.79 Sigma is 10.91 Rank: 1
High technical skills Tasks Tech index = 4.5	 Mu is 44.54 Sigma is 7.68 Rank: 6	 Mu is 35.86 Sigma is 5.31 Rank: 4	 Mu is 36.41 Sigma is 2.76 Rank: 5	 Mu is 34.31 Sigma = 2.75 Rank: 1 (not statistically significant vs.	 Mu is 34.51 Sigma is 3.63 Rank: 2	 Mu is 34.76 Sigma is 7.33 Rank: 3
High technical skills + Diverse tasks Tech index = 4.5	 Mu is 48.88 Sigma is 12.81 Rank: 6	 Mu is 39.47 Sigma is 8.17 Rank: 4	 Mu is 37.15 Sigma is 3.03 Rank: 2	 Mu is 35.54 Sigma is 2.61 Rank: 1	 Mu is 38.88 Sigma is 6.73 Rank: 3	 Mu is 47.24 Sigma is 12.18 Rank: 5

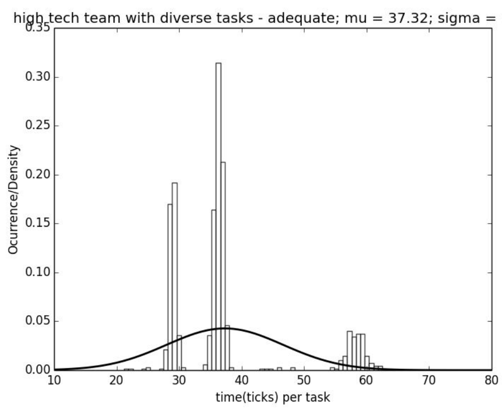
Adequate Solutions - 80% of time

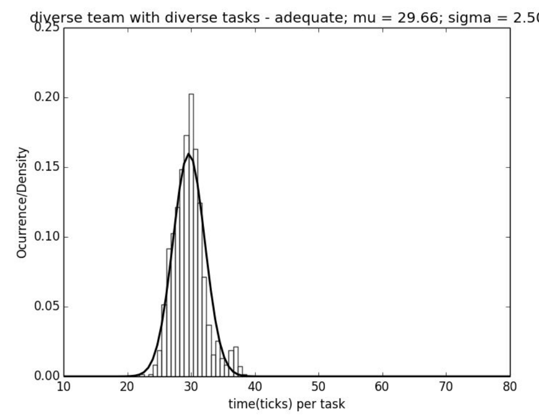
We also tested that whether lowering the solution standard, as in requiring only 80% of the full time needed to finish a task, will shift how effective diversity or tech ability at influencing the team performance. Comparing the figures below with those in the table above, we found that even though the distance between the high tech team and diverse team in terms of mean of performance time will shrink when we lower the ideal requirement of the tasks, the high tech team will never catch up with diverse team or avoid uncommon but drastically erroneous performance no matter how we lower the standard.





Ad - diverse team with high tech and diverse tasks; mu = 30.02; sigma = 2.5 Ad -high tech team with diverse tasks; mu = 37.32; sigma
= 9.34





Ad - high tech team with diverse tasks; mu = 37.32; sigma = 9.34 Ad - diverse team with diverse tasks; mu = 29.66;
sigma = 2.50

DISCUSSION

Teams perform well in tasks that correspond to their team types:
Random team performs the best in random tasks. Teams that are diverse (e.g. diverse team; high tech and diverse team; high tech and partially diverse team) perform well in diverse tasks, while teams that have high tech skills (e.g. high tech team; high tech and diverse team; high tech and partially diverse team) perform well in high tech tasks. However, high tech team is still not an optimal team structure even with high tech tasks; both high tech and diverse team and high tech and partially diverse team can outperform it. This is true even if we added communication cost for diversity.

Diversity Stabilized Team Performance:
Diversity is correlated with the appearance of “spikes” in the distribution of time. The more diverse a team is, its distribution of time has lower variance and less separate spikes. The advantages of diverse team is its stable performance across all different task types. Even when diverse team has to deal with high tech tasks, its performance is better (in terms of shorter distance to high tech team’s performance) than that of high tech team when it has to deal with diverse tasks. Therefore, the overall work efficiency of diverse teams outperform high tech teams. In fact, because high tech teams perform too poorly on any tasks involve diversity, even random teams can outperform high tech teams.

High Technical Team Tend to “Underperform” from Expectations:
From normal distribution parameters we can see that high technical abilities team (referred as high-tech team below) is more likely to have variance in their performance (greater sigma) and worst case performance, which means they are more likely to screw up. Surprisingly, the “worst-case performance” still hold for high-tech team even when they are working with high tech tasks.

From the graph we can see, no matter what task the high-tech team takes on, its performance has a small lump at the right tail, indicating a persistent phenomena of performing below expectations. The intuition behind is simple: since we assumed most people who have great technical skills developed homogeneous perspectives or roles, when they stuck in a task that needs perspectives or roles beside their owns, they are less flexible and will take a long time to get out. This further legitimizes the possibility of Apollo Syndrome (Belbin, 1981) that a group of smart people might perform surprisingly worse.

Lowering the Standard of Solution Cannot Save High Tech Team
Noticing that high tech team doesn’t do well on the right tail, we try to see whether lowering the standard of a solution (in this case, represented as spending 80% of the time as required in ideal situation). After testing in two edge cases in which the high tech team is behind the diverse team, we can see the lowering the standard can indeed shrink the distance between the diverse team and the high tech team, but it’s impossible for the high tech team to catch up or even outperform diverse team by just lowering the solution standard. The intuition behind is that since only the values have shrunk proportionally, the shape of distribution will stay the same. Both the mean and the std will decreases for the adequate (80%) solutions.

In this case, teams that have big mean values or high variance (e.g. high tech team) will benefit the most and thus have smaller mean distances to other teams, but cannot exceed the limit of proportion.

A Band-aid Diversity Solution is not Enough

In this model, we tested two types of combination of "ideal" diversity and technical abilities (high tech and diverse team & high tech and partially diverse team) and they both had shown better overall performance than diverse team or high tech team. This proves that diversity and technical abilities are not necessarily a zero sum game. They also represent two different kinds of attitudes toward diversity. The former recruiting strategy is to organize a group portfolio that for each person it selects, it considers other members in the team and whether the newcomer have something unique to offer. The latter is to prioritize technical abilities and "remedy" the team diversity through making the last person the most complementary to the rest of the team. The graphs show that high tech and diverse team have better performance over high tech and partially diverse team across all task types including high-tech abilities tasks which there isn't a statistically significant advantages for the latter ($p=0.1878 > 0.05$). This means that to fully unleash the potential of team performance, a "band-aid" diversity solution (as implied by the high-tech and partially diverse team) is not enough. We need to actively approach diversity by considering its impact for the team in every recruiting step.

REFERENCE

Belbin, R. (1981). *Management teams*. New York: Wiley.

Hong, L. & Page, S. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings Of The National Academy Of Sciences*, *101*(46), 16385-16389. http://dx.doi.org/10.1073/pnas.0403723101

Josh, H. (2014). *Mathematicians refute oft-cited 'diversity trumps ability' study - The College Fix*. *The College Fix*. Retrieved 7 October 2016, from http://www.thecollegefix.com/post/20375/

Thompson, A. (2014). Does Diversity Trump Ability?. *Notices Amer. Math. Soc.*, *61*(09), 1024. http://dx.doi.org/10.1090/noti1163

APPENDIX

Team type	**Recruiting Strategy**
Random	Agents are selected randomly
High technical abilities	For each member of the team, randomly select five agents and pick the one that has the highest technical abilities score among them.

Diverse	For each member of the team, select agents that have either different perspectives *or* roles with people already in the team.
High technical abilities and Diverse	For each member in the team, select agents that have 1) technical skills >=3 *and* 2) either *different* perspectives *or* roles with people already in the team
Low technical abilities and not Diverse	For each member in the team, select agents that have 1) technical skills <=3 *and* 2) either *similar* perspectives *or* roles with people already in the team.
High technical abilities and Partially diverse	Select the first four members based on high technical abilities, but select the last member based on diversity (who has to have different perspectives *and* roles than everyone else in the team).