IMPACTS OF INTRATEAM DIVERSITY ON TEAM PERFORMANCE VARIANCE: TWO WAYS TO TAKE CHANCES

FABRICE CAVARRETTA [1]
Fabrice.Cavarretta@insead.edu

INSEAD
Boulevard de Constance
77305 Fontainebleau
France

ABSTRACT

I predict a curvilinear relationship of diversity to performance variability, therefore risk, which I verify by reanalyzing a business simulation with 35 teams. Furthermore, combining mean and variability effects leads to improved predictions: under certain conditions, the conclusions drawn from a classical mean analysis can even reverse when considering variability effect.

LITERATURE REVIEW

Most organizational research on intra-team diversity and team performance has sought to answer the question whether diversity increases performance. However, the results have not been consistent across studies (O'Reilly, Williams, and Barsade, 1998) and puzzled researchers for decades with two views, one optimistic and one pessimistic, which directly contradict each other (Mannix and Neale, 2005). On the one hand, intra-team diversity brings more knowledge to the team (Nemeth, 1986), improving performance. On the other hand, diversity disrupts social integration (O'Reilly, Caldwell, and Barnett, 1989), impairing performance. Such approaches assume the relevant dependent variable to be the mean performance, and that the effects of information processing and social integration can neutralize each other. Nemeth and Staw (1989) demonstrated that these effects counter each other, which explains why most meta-analyses find a neutral effect (Williams and O'Reilly, 1998).

Scholars seeking to reconcile these findings have pursued two main strategies (van Knippenberg and Schippers, 2007). In the first, diversity is defined with increasing subtlety (e.g., Bunderson and Sutcliffe, 2002), leading to comprehensive theorizations of the definition of the construct (Harrison and Klein, 2007). In the second, various variables are introduced to moderate the relationship between diversity and performance (e.g., Polzer, Milton, and Swann, 2002).

However, there are contexts in which mean performance is not the performance measure of interest: in so-called “top-score competitive system” (Miner, Haunschild, and Schwab, 2003), one cares about a ranking more than about average performance. At the extreme, mean

[1] I am grateful to Martin Kilduff, Reinhard Angelmar and Ajay Mehra for sharing their dataset. I thank my thesis adviser, Bruce Kogut, for his support, as well as Rodolphe Durand, Henrich Greve, Etty Jehn, Denrell Jerkell, Christoph Loch, Randall Peterson, and Maurizio Zollo for their thoughtful comments on earlier drafts.

performance is uninteresting in a winner-take-all contest, where one would want to know what are the factors leading to a top-position. For that purpose, performance variability—also called risk—may matter as much as does mean performance in organizational research (March, 1991; Baum and McKelvey, 2006). For instance, an influencing factor may have a negative effect on the mean performance but a positive effect on variability. Then, an increase in the factor may increase the probabilities to reach the top rank in a competition as long as the variability effect is strong enough.

Most research on diversity explores its mean effect using linear regressions. However, this method does not adequately address the question of whether it increases the probability of finishing with either very high performance or very low performance. The current study focuses on the effect of diversity on performance variability (Fleming, 2004; Taylor and Greve, 2006).

THEORY

Many studies view diversity as a variable with constant effects at any level. Yet, teams with high diversity have the potential for low integration and high knowledge, and they differ fundamentally from homogeneous teams that have a potential for low knowledge and high integration. Hence, I distinguish between mechanisms operating at *high* and *low* levels of diversity, similar to theorizing that hypothesizes curvilinear effects of diversity (e.g. Gibson and Vermeulen, 2003; Richard, et al., 2004; Uzzi and Spiro, 2005).

In high diversity levels (low-integration and high-knowledge teams), greater diversity leads to a higher probability that team members do not get along which thus reduces the performance of the lowest performing teams. However, for the teams where members do get along, greater diversity increases knowledge, thus improving performance of the best performing ones. It implies that within a population of high diversity teams, some teams fail and some succeed, and the spread is exacerbated by greater diversity. In other words, in addition to possible effects on average performance, greater diversity leads to an increase in performance variability. This conclusion matches previous claims (Taylor and Greve, 2006), except that instead of predicting it applies constantly, I expect it to occur only for high levels of diversity.

Hypothesis 1: In a population of teams with high intra-team diversity, greater diversity leads to higher variability of team performance.

In low diversity levels (low-knowledge and high-integration teams), greater diversity improves available knowledge and therefore leads to a greater probability that the team does not lack knowledge crucial to accomplish its task—thus improving performance of the lowest performing teams. At the same time, for the teams that possess the proper knowledge, greater diversity implies a reduction of social integration—thus lowering performance of the highest performing teams. It implies that in a population of low diversity teams, some teams fail and some succeed, but the spread is reduced by greater diversity. In other words, greater diversity in low levels of diversity may not change the expected performance but results in a reduction of performance variability.

Hypothesis 2: In a population of teams with low intra-team diversity, greater diversity leads to lower variability of team performance.

METHOD

Data

The empirical analysis was designed to test the hypotheses on variability, as well as comparing mean vs. variability effects of diversity. To emphasize the counter-intuitive nature of the conclusions, I therefore chose to reanalyze the data from an existing study of diversity (Kilduff, et al., 2000) that its authors accepted to share with me. The research setting was a MARKSTRAT business simulation, a game in which groups of players compete as management teams in a simulated market comprising five competitors. The game ran for six periods during which teams made decisions on marketing, production, and R&D, and a computer determined competitive performance results by market share and revenues. The sample consisted of 159 business executives divided into 35 teams, which were the unit of analysis.

Variables

Independent Variables. I used three demographic variables (age, nationality, function), which were computed by a coefficient of variation or Blau's index of heterogeneity. I added two cognitive diversity variables, captured through items in the questionnaire on items derived from questions used by Zucker (1977) to measure cognitive variability in an institutionalization process. The first cognitive variable, diversity in s*pecialization perception*, was measured by asking participants how specialized the team members were, ranging on a scale from "no person has a specialized role to play" to "each person has a specialized role to play." The second cognitive variable, diversity in *power perception*, was measured by asking participants how easy it would be to challenge the decision-making power of the dominant members, ranging on a scale from "very easy to challenge the decision-making power of the dominant members" to "very hard to challenge decision-making power of the dominant members."

Determining diversity level of the teams' population. The empirical setting presents a natural experiment on the demographic variables, and an empirically clear situation on the cognitive variables. The sample of teams is either high or low in diversity, for each variable, which allows for testing a linear hypothesis on each variable (either H1 or H2); this setting is convenient since the small sample size (N=35) would not allow testing full curvilinearity. The variables for which all the teams low diversity (age, nationality, perception of power and specialization) allow testing of hypothesis H2, and functional diversity, for which the teams high diversity, testing of hypothesis H1. The effect of diversity is assumed locally linear, reducing (H2) or increasing (H1) performance variability, respectively.

Control variables. Since the current study builds on an existing mean effect analysis (Kilduff, et al., 2000), I keep the control variables of the previous study. Therefore, the size of the team and a measure of starting position—each team in the simulation did not start with the same market share—is included. Furthermore, all variables identified in the original study as relevant independent variables, but not used here as independent variables, appear as control variables. These include the diversity in perception of *ambiguity*, *decision difficulty*, *decision pressure* and *effectiveness* (see Kilduff, et al., 2000 for more details on those variables).

Dependent Variables. Two performance measures—*Cumulated Net Marketing Contribution* (CNMC) and *Final Market Share* (FMS), expressed as a percentage—reflect

outcomes. These variables are the dependent variables of the mean analysis, and their variances are the dependent variable of the variability analysis.

Analyses

I first perform an OLS multiple regression for both performance measures to determine the mean effects. A first approach to detecting an effect of a variable on the variance assumes a parametric form of the absolute value residual of the regression and test for it (similar to Taylor and Greve, 2006:732). Because of possible bias on that method, I also use the Goldfeld-Quandt test, designed to determine the direction of heteroskedasticity, even with a small sample.

To determine when the variability effect overcomes the mean effect, I consider performance in each sub-sample is a random variable with a mean and a variance and compute the performance level Y_c where the cumulated probabilities equalize. For values around the average of performance, the effect is dictated by the mean effect estimated by the regression; beyond that critical level of performance, the effect reverse.

Results

Hypothesis testing. Hypothesis 1 predicts that functional diversity should increase the variance of performance. This result should appear in the regression as a positive coefficient for that diversity variable, and a ratio of variance greater than one in the Goldfeld-Quandt. The effect is in the right direction for all four measures, and significant for the regression of the absolute residual of FMS, so hypothesis H1 is supported. Hypothesis 2 predicts that diversity in nationality, age, perception of specialization and power should decrease variance since the teams' population was at low diversity levels for those variables. This result should appear in the regression as a negative coefficient for those diversity variables, and in the Goldfeld-Quandt test as a ratio lower than one. The results show a pattern of confirmation for all four diversity variables so, hypothesis H2 is supported.

Combining mean with variability effects. I then conduct an exploratory analysis examining whether effects of diversity on performance variability better predicts extreme performance. Since the teams were in low age diversity levels, an increase in age diversity reduces performance variability. I show a counter-intuitive effect: low age diversity is associated with greater success probability—than high age diversity—when the performance goal is to belong to the best 18% of the teams. Therefore, if a team aims not for average performance but for a high performance threshold, then lower age diversity should be preferable to higher diversity. This conclusion nuances the conclusion drawn from an initial analysis on the same data set by Kilduff, Angelmar and Mehra (2000): they found that greater age diversity increases mean performance, implying that age diversity is therefore preferable to lower age diversity. Hence, although age diversity may increase mean performance, it reduces the probability of achieving high performance.

DISCUSSION

This study intends to make two contributions. First, it shows that demographic and cognitive diversity influences not only mean team performance but also has a U-shaped influence on the performance variability, or risk. Second, it shows, by using a quantifiable

criterion, that the variability effects lead to more nuanced conclusions about diversity than by just using a mean effect.

Resolving contradicting theories of diversity

Two contradictions occur among theories of diversity. Those considering the effect of diversity on performance struggle between a pessimistic perspective—diversity decreases social integration, therefore performance—and an optimistic perspective—diversity improves available knowledge, therefore performance. Such a contradiction underlies the non-significant effect of team diversity on team performance. However, introducing performance variability allows us to develop hypotheses that predict both outstandingly good and bad outcomes even in the absence of mean effect. In the empirical analysis, four diversity variables had a neutral effect on performance yet had a significant effect on risk.

A second contradiction appeared in previous studies considering the effect of diversity on performance variability. A knowledge perspective suggests that diversity creates risk (Fleming, 2001; Fleming, 2004; Taylor and Greve, 2006), while a socio-psychological perspective suggests that similarity creates risk (Schachter, et al., 1951; Janis, 1971; Luthans, 2002). Here, I propose a curvilinear relationship, whereby risk appears at both extremes—high and low—of the diversity scales.

This approach therefore distinguishes taking chances on the internal dynamics when in high diversity, from taking chances on the task-knowledge match when in low diversity. It predicts both exceptionally high and exceptionally low outcomes for teams with both highly diverse and highly similar members, and bounded outcomes for mid-range diverse teams.

Applications to theories of organization

The variability analysis proposed above applies to contexts that reward or punish extreme outcomes, therefore risk, and could generally be applied to any context where a variability effect occurs and the organizational goal can be expressed as reaching a threshold. This resonates with the call to arms voiced by some scholars to expand our focus beyond effects on the mean (Cavarretta, 2007). For instance, Starbuck (1993) claims organizational studies should focus on exceptional organizational events and praises the explicit analysis of outliers in an approach that requires reconsideration of assumptions about the distribution of outcomes. Daft and Lewin even recommend the exploration of "heretical methods" by proposing the preliminary study of outliers as a potential way of renewing organizational studies (1990).

When extremely low outcomes matter. First, consider fields that punish going below a minimum performance threshold. For example, the field of High-Reliability Organizations (HRO) studies organizations dealing with low and disastrous outcomes. For instance, Perrow (1984) studies nuclear power plants, where reaching a low threshold can lead to a devastating accident. Weick (1990; 1993; Weick and Sutcliffe, 2001) explores the factors leading to low performance such as a plane crash, or human losses in firefighting. In these HRO contexts, both researchers and practitioners care about the impact of organizational factors—such as demographic and cognitive diversity—on risk. The current paper shows how diversity might have no mean effect and yet influence risk, therefore the chances of catastrophic outcomes.

When extremely high outcomes matter. Other contexts force organizations to take risks by rewarding those reaching high thresholds of performance, or by letting only a few

winners of a contest to survive. In entrepreneurial settings, reaching high and rare performance levels (such as profitability, network effects, IPO, etc.) leads to survival, suggesting that effects on variability may matter. Having a greater chance of ranking at the top of one's cohort, when trying to reach IPO stages, for instance, may be more predictive of final success than merely improving average performance. Because of the existence of thresholds of survival, entrepreneurs or venture capitalists may benefit from effects on variance by manipulating diversity and consensus in their teams.

CONCLUSION

Organizational scholars have long suspected that intra-team diversity correlates to extreme organizational outcomes such as fiascos (Janis, 1971). Proving the relationship has remained elusive over the years (Nemeth and Staw, 1989), even though a rich stream of literature has been written on the subtle effects and varied dimensions of diversity. Focusing on the performance variability (Sorensen, 2002; Denrell, 2003; Taylor and Greve, 2006) offers an alternative approach to studying the effect of team diversity. Considering both information processing and social integration, I propose that in a population of teams, pushing toward the extremes of intra-team diversity increases inter-team performance variability. Furthermore, such variability amounts to risk effects—increasing the chances of reaching or missing thresholds of performance—that can run in opposite directions of the mean effects. Under reasonable assumptions, conclusions drawn from a conventional mean analysis—e.g. age diversity improves mean performance—can even reverse—e.g. age diversity is *detrimental* if trying to reach top 10% of the population—when considering variability effect. In other cases, the existence of a variability effect helps predict outcomes, even though no significant mean effect of diversity exists.

The paper shows how studying the variability effects of diversity may explain extreme organizational outcomes. Such logic informs any area of organizational study where reaching or avoiding a certain performance threshold matters most. Low thresholds of performance matter to HROs and governance literature, since both explore the incidence of major accidents, bankruptcy, or fiascos. It suggests that cognitive and demographic diversity influence the chances of organizational accidents, beyond the mean effects of those factors. In entrepreneurship and population ecology approaches, high thresholds of performance matter, therefore effects on performance variability leading to high and rare outcome may be more relevant than a mean effect on performance. In such contexts, cognitive and demographic diversity might then increase the chances of survival, even when these factors degrade mean performance.

The long tradition of studying the effect of diversity on performance has been frustrated by the contradictory effects of underlying mechanisms. In the meantime, practitioners have not waited for the conclusion on that matter, and currently push to develop diversity at any cost. In a world where both risk—industrial, governance, entrepreneurial, etc.—positively or negatively matter to organizations, establishing that team diversity grows risk at both extremes of its scale should warrant further study.

REFERENCES AVAILABLE FROM THE AUTHOR