DECISION SCIENCES
A JOURNAL OF THE DECISION SCIENCES INSTITUTE

Decision Sciences
Volume 00 Number 0
xxxx 2018

Shared Leadership Regulates Operational Team Performance in the Presence of Extreme Decisional Consensus/Conflict: Evidences from Business Process Reengineering

Manfredi Bruccoleri[†]
Department of Industrial and Digital Innovation, University of Palermo, Viale delle Scienze Ed.8, Palermo 90128, Italy, e-mail: manfredi.bruccoleri@unipa.it

Francesca Riccobono
Department of Industrial and Digital Innovation, University of Palermo, Viale delle Scienze Ed.8, Palermo 90128, Italy, e-mail: francesca.riccobono@unipa.it

Andreas Größler
Institute of Business Administration, Operations Management Department University of Stuttgart, Keplerstraße 17, Room 10.037, 70174 Stuttgart, Germany, e-mail: andreas.groessler@bwi.uni-stuttgart.de

ABSTRACT

This study focuses on decision-making within operational teams. Grounding our argumentation on group decision-making literature, we argue that adverse behavior patterns may affect the way in which consensus is achieved within the team, and that team performance has an inverted U-shaped relationship with the level of consensus. Then, by relying on leadership literature, we pose the hypothesis that the level of shared leadership inside the group moderates this U-shaped relationship. To empirically test our literature-based argumentation, we use longitudinal data collected in the years 2014 and 2015 from Business process reengineering projects, each lasting three months, conducted by 141 Master of Science Students grouped in 34 teams. We conclude by emphasizing that it is important to control for the occurrence of behaviors which lead to "fake" consensus within operational teams, by observing the individuals' satisfaction with respect to the group decision as well as their active participation in the decision-making process.

[Submitted: March 10, 2017. Revised: April 27, 2018. Accepted: June 12, 2018.]

Subject Areas: Abilene Paradox, Business Process Reengineering, Group Decision-Making, Groupthink, Shared Leadership, Operational Team, and and Performance.

[†]Corresponding Author

INTRODUCTION

As long as organizations deal with increasing demands for efficiency and responsiveness, team-based work structures have become a pervasive organizational model to face such challenges (Boyett & Conn, 1992). Relying on effective and efficient groups at all organization's levels is fundamental in order to achieve sustainable competitive advantage. In fact, not just top management teams, but also operational teams play a crucial role in determining processes' efficiency, customer satisfaction, and company success (Bamford & Griffin, 2008).

The effectiveness of a group decision-making process may be, however, threatened by a few psychological behaviors which influence the way groups make the decision (Riccobono, Bruccoleri, & Größler, 2016; Swaab, Phillips, & Schaerer, 2016). Two organizational behavior theories exist underlining the potential pitfalls of a group decision-making process, but they explain two different phenomena: the "Abilene Paradox (AP)" (Harvey, 1974), and "Groupthink (GT)" (Janis & Mann, 1977). The extent to which such phenomena could be a real threat in an operational context has been described by Riccobono et al. (2016) and an example of how they lead groups to ineffective behaviors is reported by McAvoy and Butler (2009), in their empirical study on two software development teams trying to reach consensus on the way a software package must be developed. In one case, the group chose to adopt a prioritization method different from the one recommended by the Agile Software Development approach, and, even if it was proven to be ineffective, the group seeking for unanimous consensus persisted in adopting it. Here the authors identify GT behavior, where the team achieves consensus mostly because of their "seeking for unanimity" willingness. In the other case, the group, even though not all group members agreed, decided to follow standardized company procedures in the adoption of the Agile approach that were not fitting the specific project, and that finally led to the failure of the project. Here the authors recognize the presence of the AP behavior, where the group decision is made, even with little consensus, because of the team members' "conforming to the others" inclination.

When the group experiences the AP, members are usually poorly satisfied with the final group decision, also given the high level of divergence between the decision and their thoughts: the resulting final decision is likely to be inefficient. Contrarily, in the case of GT, poor group decisions are usually associated with a high level of satisfaction experienced by group members, often associated with a low level of divergence between their opinion and the final group decision (Sims, 1994). The AP and GT may thus be differentiated by the level of *group decision consensus* (GDC), meant as the members' satisfaction about the group decision for a given level of divergence between each member's opinion and the final group decision. A group's final decision which is associated with a very low or a very high level of consensus, is likely to be ineffective and, in turn, to generate poor group performance (Schweiger, Sandberg, & Ragan, 1986; De Dreu & Weingart, 2003). While decision making research has largely focused on how top management groups reach consensus (e.g., Schweiger et al., 1986) or experience conflict (e.g., Jehn & Mannix, 2001) and on resulting group performance implications, few studies investigate group decision-making consensus in operational groups.

Moreover, the psychological and the organizational behavior literature underline the key role played by leadership for enabling group decision-making efficiency (Hackman, Walton, & Goodman, 1986; Cohen & Bailey, 1997; Alby & Zucchermaglio, 2006). Some researchers have even argued that the leadership provided by team members, referred to as *shared leadership*, is the most crucial ingredient (Sinclair, 1992; Stewart & Manz, 1995; Zaccaro, Rittman, & Marks, 2002; Kozlowski & Bell, 2003). However, most of the existing research on operational teams' leadership has focused narrowly on the influence of vertical leadership, that is, a manager external to a team having formal authority over it (e.g., McFadden, Henagan, & Gowen, 2009), while neglecting the role of shared leadership, especially in relation to undesirable group behavior such as the AP and GT.

Our study addresses these gaps and wishes to answer the following research questions: are operational teams subject to adverse decision-making dynamics similar to top management teams? Is it important to reach a certain level of consensus when making operational decisions and how does this influence group performance? What is the role played by shared leadership in dampening the negative effects of too little (Abilene-like) and too much (GT-like) consensus?

To answer these questions, we develop a conceptual model and two hypotheses based on the theoretical underpinnings of the AP and the GT phenomenon. We provide theoretical arguments about the implications of reaching consensus on the effectiveness of the decisions undertaken by groups. Also, we test the moderating effect of the level of shared leadership within the group on the relationship between the level of consensus reached and the effectiveness of the decision. The hypotheses are tested in an experimental research setting consisting of 34 three-month long business process reengineering (BPR) projects. We use primary data from 377 individual-level observations collected from 141 students at three key decisional time points of the projects' life cycle. These are then aggregated into 100 group-level observations related to 34 groups involved in the projects.

In addition to contributing to the literature on organizational behavior, operations management, decision science, and leadership, this article offers managerial suggestions for operational group managers who experience GT-like and AP-like consensus behaviors.

LITERATURE REVIEW

In what follows we provide the theoretical background of the two behavioral group decision making phenomena, namely GT and the AP, and of the shared leadership concept, that lay the foundation for our hypotheses development. In particular, we describe their origins, meanings and the streams of literature which have been developed around them, with a focus on studies conducted in operational contexts. Then, we explain how our study is positioned within these fields of studies.

GT Behavior

Janis (1982, p.9) describes GT as a "mode of thinking that people engage in when they are deeply involved in a cohesive group, when the members striving for unanimity override their motivation to realistically appraise alternative courses of action." The occurrence of GT has been initially studied in political and military contexts where decision-making processes led to fiascos (e.g., the Japanese Pearl Harbor attack to U.S. fleet in 1941). In fact, Irving Janis, the founder of this theory, was inspired by being an active participant in the U.S. army during the Second World War.

Successively, the GT phenomenon has been investigated in business settings. However, results from empirical research provide only partial validation for it (Leana, 1985; Moorhead & Montanari, 1986; Herek, Janis, & Huth, 1987; Esser, 1995, 1998). The controversial results, subsequently, led the literature to split into two main fields of research: one that completely reconceptualizes the basic theory on GT behavior, its antecedent conditions and consequent effects on decision-making behavior in groups (Neck & Moorhead, 1995); the other one that, starting from the original GT model (Janis & Mann, 1977), looks at new variables (e.g., Neck & Moorhead, 1995) and proposes different linkages among the constructs (Park, 2000; Chen, Tsai, & Shu, 2009). Finally, GT has been empirically disclosed in various operational contexts, such as production and quality control work groups (Manz & Sims, 1982), new product development teams (Brockman, Rawlston, Jones, & Halstead, 2010), and in disaster operations management.

The AP Behavior

A related cause of ineffective decision-making is referred to as the AP. Harvey (1974) originally described AP as instances when group members, and hence groups as a whole, "frequently take actions in contradiction to what they really want to do" (p. 18) and this behavior reflects the group's "inability to manage agreement" (p. 65). The author describes AP through a parable about four family members on a porch in Coleman, Texas, on a very hot summer day. While everyone appears happy drinking lemonade and playing dominoes, someone in the group suggests taking a drive to Abilene (about 53 miles from Coleman) to eat lunch. Individually, each of the four persons thinks this suggestion was without merit because the only available car has no air-conditioner. But each one goes along, so as not to be perceived as a "spoiler" of the group. Upon returning exhausted and disgruntled, the family members recognize that not even one of them really wanted to make the trip. They were unable to justify their original decision to take a 106-mile drive in a dust storm merely to eat a mediocre lunch in such hot weather. In the organizational realm, Harvey provided an example of how AP evolved in a company where top management wanted to go with its "favorite" project to Abilene, while all the individual members of the research and development (R&D) department wanted to go somewhere else.

Contrary to the notable literature developed on GT, interest on AP was substantially lower, overshadowing managerial attention to AP (Harvey, Novicevic, Buckley, & Halbesleben, 2004). Very few studies exist empirically analyzing this phenomenon, and of these the most are applied in the field of software development projects (Appan, Mellarkod, & Browne, 2005). For example, McAvoy and Butler (2009) conduct a longitudinal case-study over two project teams (both of which comprised six developers and a project manager). They find the existence of GT in one case and of AP in the other case. For example, regarding the GT symptom *rationalization to delete feedback and views opposite to group position*, the authors report "When it became clear that the decision to change the prioritization method recommended by Agile approach was a bad decision, the correct method was not reconsidered for use" (McAvoy & Butler, 2009, p. 377). In the second case study, regarding the AP symptom *organization members fail to accurately communicate their desires and/or beliefs to one another*, the authors report that "Although initially and subsequently agreeing with the decision to adopt Agile, developers did not express opinions when Agile use gradually died" (McAvoy & Butler, 2009, p. 380).

Differences between GT and AP

Few authors have attempted to distinguish the behavioral mechanisms associated with GT and AP. According to Sims (1994), AP is similar to GT, but also differs in significant ways, including that in GT individuals are not acting contrary to their conscious wishes and generally feel good about the decisions the group has reached. Instead, in AP, individuals acting contrary to their own wishes are more likely to have negative feelings about the outcome. In other words, GT is a psychological phenomenon affecting clarity of thought, whereas in AP clarity of thought is unaffected.

Harvey et al. (2004) adapt the approach proposed by Kim (2001) and present a comparison between AP and GT. In their comparison, the main differences among the two phenomena are related to the following points: (i) cohesiveness among group members is very low in a situation characterized by AP while very high in a situation showing GT; (ii) leadership is ineffective in AP while overpowering in GT; (iii) the perception of different points of view is not salient in AP while perceived as an enemy in GT; (iv) the group members' points of view are private-oriented in AP while group-oriented in GT; (v) the feeling of group members with respect to the group decision is a feeling of nonresponsibility in AP versus full awareness in GT; (vi) the level of satisfaction about the group decision is very low in AP while very high in GT; (vii) the blame in case of group failures is attributed to (other) group members in case of AP while to external factors in GT.

Shared Leadership and Group Decision-Making

Carson, Tesluk, and Marrone (2007), p. 1218) define shared leadership as "an emergent team property that results from the distribution of leadership influence across multiple team members. It represents a condition of mutual influence embedded in the interaction among team members that can significantly improve team and organizational performance." According to this conceptualization, shared leadership ranges along a continuum based on the number of leadership sources (i.e., team members) having a high degree of influence in the team. Accordingly, the low-end of the continuum captures groups in which team members follow the leadership of a single individual; hence there is a single source. In contrast, at the high end of the shared leadership continuum there are teams in which most, if not all, members provide leadership influence on one another; thus, there are multiple sources of leadership distributed among team members rather than concentrated or focused on a single individual. In fact, shared leadership is proposed as a contrasting paradigm with respect to *vertical leadership* (Pearce & Sims, 2002), where the manager/leader of the group is positioned hierarchically above and, often, is external to the team, has formal authority over it, and is responsible for its processes and outcomes (e.g., Hackman et al., 1986; Druskat, & Wheeler, 2003).

In psychology and management, many empirical works exist demonstrating the links between shared leadership and top management team performance (e.g., Avolio, Jung, Murry, & Sivasbramaniam, 1996; Pearce & Sims, 2002; Sivasubramaniam, Murry, Avolio, & Jung, 2002; Ensley, Hmieleski, & Pearce, 2006). Even early leadership scholars argued the importance of leadership shared among multiple individuals inside an organization and saw leadership as a group quality (Gibb, 1954). Shared leadership can provide organizations with competitive advantage through increases in commitment, in the personal and organizational resources brought to bear on complex tasks, in openness to reciprocal influence from others, and in the sharing of information (Katz & Kahn, 1978).

Positioning of This Research within the Literature

Although most group decision-making research has investigated issues related to decision conflict and consensus in top management teams (e.g., Simons & Peterson, 2000), we already mentioned that poor group decisions can bring about dangerous implications also in an operational context. For example, as reported in Bendoly, Croson, Goncalves, and Schultz (2010a, p. 444), "groups subject to GT can implement new quality control policies that do more to complicate work than to allow for easy identification of bottle-necks and failure points. Groups subject to AP can generate new product designs that fail to incorporate a host of distinct integration capabilities necessary for mid-term market adaptability (even if these capabilities had been individually considered crucial by various members of the design group)."

While in the field of operations management only few studies investigate the impact of group decision-making behavior on performance (such as in Bendoly, Thomas, & Capra, 2010b), understanding the relationship between operational team decision process and performance is still a hot topic in project management, information systems, organization science, and the general management literature, as demonstrated by recent studies from different streams of literature that have focused their attention on factors enabling effective team decision through team debate. In the field of project management, the study of Bendoly (2014) demonstrates that in a group project context, the group members' understanding of system dynamics, as a form of methodological expertise, enhances group project performance because of its ability to increase the quality of information shared in the group setting as well as the psychological safety among team members. Sharing information that is useful to support the project work (quality of shared information) and overcoming fear in posing inquiries and expressing ideas (psychological safety) are typical behaviors of group members in the presence of shared leadership.

Also, a recent study on human resource management by O'Neill and McLarnon (2017) has underlined the importance of psychological safety for group decision conflict effectiveness. In their review on task, relationship, and process conflict, the authors develop the "Team Conflict Dynamics Model" to connect conflict profiles (healthy vs. unhealthy) and conflict management process (collectivistic vs. individualistic) with team members' psychological safety and team performance.

In information management, Dong, Zhang, and Herrera-Viedma (2016) propose a novel consensus framework for group decision support systems, based on a self-management mechanism to manage noncooperative behaviors in the consensus reaching process (CRP). In this novel consensus framework, the experts involved in the consensus process provide not only preference information about alternatives but also mutual evaluation information for experts (professional skill, cooperation, and fairness).

A recent organizational behavior study by Tsai and Bendersky (2015) focuses on the paradoxical effects of group members' opinion differences on group performance. They demonstrate that task conflicts that are expressed as debates, rather than as disagreements, are associated with greater information sharing and in turn with greater group performance.

Although the above mentioned studies underline the important role played by factors such as information sharing, psychological safety, collectivistic behavior and mutual evaluation, and some others (e.g., Simons & Peterson, 2000) analyze the contingent role played by team characteristics, such as team climate, trust and norms, in reducing or even reversing the negative relationship between task conflict and overall team effectiveness, none of them consider the role played by the mutual influence embedded in the interaction among team members (i.e., shared leadership rather than vertical leadership) in moderating such a relationship.

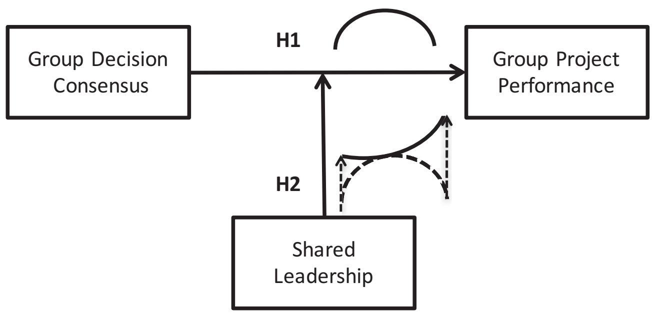
Moreover, we observed that most of these studies focus on team conflict while ignoring the potential drawback of both sides of extreme consensus. We have argued that both GT and AP concur in reaching the final group decision, but they are diverse in terms of how much consensus is reached in that decision. In this direction, the conflict management literature already provided evidence for a curvilinear relationship between task conflicts and team innovative performance, i.e., moderate levels of task conflict may promote team innovation (De Dreu, 2006). However, the relationship between operational team performance and the level of conflict/consensus remains understudied.

Our study is thus positioned within the interdisciplinary group decision-making literature and wishes to extend the understanding of the implications of GDC on operational group performance by using the lens of GT and AP as adverse behavior modes which lead the group to consensus. Also, this study is distinct from other studies on leadership in operational contexts because we investigate if and how the presence of shared leadership within the group can play a crucial role in dampening the negative consequences of GT- and AP-like consensus.

HYPOTHESES DEVELOPMENT

We develop our hypotheses about the linkages among GDC, shared leadership, and group project performance by grounding our argumentation on the theories of

Figure 1: Predicted relationships.



GT and the AP. Such linkages are represented in the conceptual model shown in Figure 1.

GDC and Project Performance

When a group makes a decision, each member feels a certain level of satisfaction regarding this decision, dependent on the level of dissimilarity between her/his personal opinion and the group's final decision. Accordingly, a certain level of consensus is associated to each group decision. The level of GDC captures the extent of group members' satisfaction, given the gap between their individual opinions and the "consented" decision.

We argue that a low GDC may reveal that behind the taken decision there is an AP-like consensus behavior; the group decision will very likely be ineffective, and this will negatively affect group performance. This line of reasoning can be explained by the psychological theories of social conformity and social influence which suggest that human beings are often very averse to acting contrary to the trend of a group (Asch, 1955; Deutsch & Gerard, 1955). According to Harvey (1974), indeed, the AP phenomenon may occur when individuals experience action-anxiety concerning a potential negative reaction and displeasure of the group if they express their personal private opinions. In this case, the groups will not exploit the benefits of social interaction, which are indeed particularly relevant when making a decision requiring creativity and problem-solving skills (Perry-Smith & Shalley, 2003).

Due to social conformity attitudes, members may even conform to the thinking of an inconsistent leader (who, in fact, "drives the group to Abilene") and the quality of the group decision suffers from a lack of confrontation and discussion between members. The group will miss the network effect coming from multiple individuals interacting with one another and bringing in their own motivation. Accordingly, members will more likely be frustrated and unsatisfied with the group decision, being so dissimilar from their private and tacit view. In sum, when GDC is low it is likely that the group decision-making process is affected by negative influences stemming from members' social conformity and influence attitudes: group members did not express their personal view, did not identify new alternatives. Still worse, they neither learnt from each other, as it is contrarily supposed to happen in typical people interaction environments according to social learning theory (Rotter, 1982). In other words, low levels of GDC might be a symptom of AP decision-making behavior which has negative effects on decision-making outcomes.

On the other side, a certain level of conflict in group decision-making has been found to provide benefits to group performance (Jehn & Mannix, 2001). The members' exchange of potentially contrasting personal opinions on relevant and critical decisions of the project can contribute to the final success of the project (Schmidt, Montoya-Weiss, & Massey, 2001; Hoegl, Parboteeah, & Munson, 2003). Such interaction, in fact, helps to reduce ambiguity and uncertainty (Dennis, Rennecker, & Hansen, 2010) related to interdependent activities and facilitates further learning among members and the group as a whole, with associated positive gains in project efficacy (Edmondson, 1999; Edmondson, Bohmer, & Pisano, 2001; Huckman, Staats, & Upton, 2009).

However, as suggested by much of the existing literature on GT, these benefits are likely to exhibit diminishing marginal returns in terms of group performance as too much—potentially irrational—consensus-making arises (Janis, 1982). In fact, a decision "shared too much," extraordinary satisfaction of group members, and quasi-inexistent differences between the individual and the group views, are likely to be symptoms of the GT phenomenon. As already mentioned in the literature review section, while in AP the individuals conform to the emerging group decision even if their private opinion remains different until the end of the decision-making process and in turn they experience negative feelings about the outcome, in GT individuals' goals and perspectives change to reflect the desire to conform to the group and they generally feel good about the decisions the group has reached (Sims, 1994). Thus, while we suppose that group performance increases with a higher level of consensus due to a higher level of group inter-personal interaction (Mohammed & Ringseis, 2001; Jong & Ruyter, 2004), we also expect that, beyond a certain level of GDC, negative GT dynamics may outweigh positive group interaction dynamics, resulting in negative project group performance. Specifically, high levels of GDC may signal a trend of group members to conform to the group and accordingly the contributions of each individual in the decision-making process decreases: the group performs worse.

Given these distinct and countervailing forces associated with GDC, we expect that group project performance will decrease as GDC decreases from a moderate to a low level (i.e., the group "is going to Abilene"), and as GDC increases from a moderate to a high level (i.e., the individuals are "thinking too much like the group"). Therefore, we propose the following hypothesis.

H1: Group project performance has an inverse U-shaped relationship with level of GDC.

The Moderating Effect of Shared Leadership

As already mentioned in the literature review section, shared leadership captures the distribution of leadership influence across team members. We argue that shared leadership moderates the relationship between GDC and group project performance.

Specifically, when the level of shared leadership within the group is low, it is more likely that the negative effects of having too low or too high GDC will be exacerbated. In fact, when a decision is reached with a low level of GDC, the likelihood that this is a symptom of the AP phenomenon increases. The group members will miss to exercise positive influence among each other because they do not try to influence each other by expressing their personal view supported by their own motivation, while allowing an inconsistent leader to drive the group to Abilene. They hesitate to exchange information among each other and will not gain the benefits advocated by social network theory (Sparrowe, Liden, Wayne, & Kraimer, 2001). Analogously, holding shared leadership low, when the group members show a very high level of consensus regarding the decision-making outcome, the likelihood that this is a symptom of GT phenomenon increases. Group members will more likely change their perspectives to reflect the desire to conform to the group, will allow a solitary and directive leader to guide the group towards an irrational and fake "shared" decision, and will not experience valuable interactions and constructive discussion.

Contrarily, in groups where the level of shared leadership is high, ties between team members are stronger because they perceive each other as exerting leadership influence in the team. Group members not only exchange their opinion but also experience deeper and more effective communication among them. Accordingly, in this case, during the decision-making process group members develop virtuous ideas exchange mechanisms, that in turn lower the likelihood of negative group dynamics phenomena (GT and AP) to occur. More specifically, in this case, if the group members show a high level of GDC at the end of the decision-making process, this will more likely be a signal that the final group decision was really "shared"; if they show a low level of GDC, this will more likely be neither a signal of "inability to manage agreement" nor a signal of hiding personal points of view just to avoid being perceived as "spoilers" of the group, but rather this will be a signal of heated argument and passionate discussion, which can increase the quality of the group decision even if it does not lead to consensus.

Taken together, the above arguments suggest that the level of shared leadership within a group moderates the relationship between GDC and group project performance, because: (i) group decision-making processes benefit from valuable group exchanges of opinion and information; (ii) high consensus can be interpreted as a signal of a "really joint" decision in groups with high level of shared leadership, while they signal "fake joint" decisions in groups with low level of shared leadership; (iii) in groups with high level of shared leadership, low consensus can be interpreted as the result of a discussion which was really animated and examined in depth by members with committed divergent opinions, while in groups with low level of shared leadership, low consensus is the result of members' failure to accurately communicate their desires and/or beliefs to one another.

Accordingly, we expect the inverse U-shaped curve between GDC and group project performance to flatten its shape until reversing it. In line with these arguments, we propose the following hypothesis:

H2: In group decision-making, the level of shared-leadership within a group moderates (flattens or reverses) the inverse U-shaped relationship between group project performance and level of GDC.

RESEARCH DESIGN

Data Collection and Sample

We conducted two longitudinal field experiments, each across a three-month time frame. The first experiment ran from October to December 2014, the second from October to December 2015. The experiments involved in total 141 first-year MSc students (59.7% male, 24 years old on average) of two different academic years (i.e., 2014–2015 and 2015–2016) that had to carry out a BPR project in a real company, as a part of the business process management (BPM) 9-ECTS class, at a large Italian university. The participating students were grouped into 34 groups, of which 27 made of four members each, 6 groups of five members each, and one group of three members. The three-month BPR projects constituted our research setting. A BPR project deals with the fundamental rethinking and radical redesign of business processes to achieve dramatic improvements in critical measures of performance, such as cost, quality, service, and speed (Hammer & Champy, 1993). The core phases of a BPR project are mapping and analyzing the "AS-IS" configuration of a given process within the organization, and designing and testing the "TO-BE" configuration, which can overlap the main criticalities, previously identified, and in turn improve the organization's business performance. Specifically, the BPR projects of the BPM class considered in our study are divided into four phases.

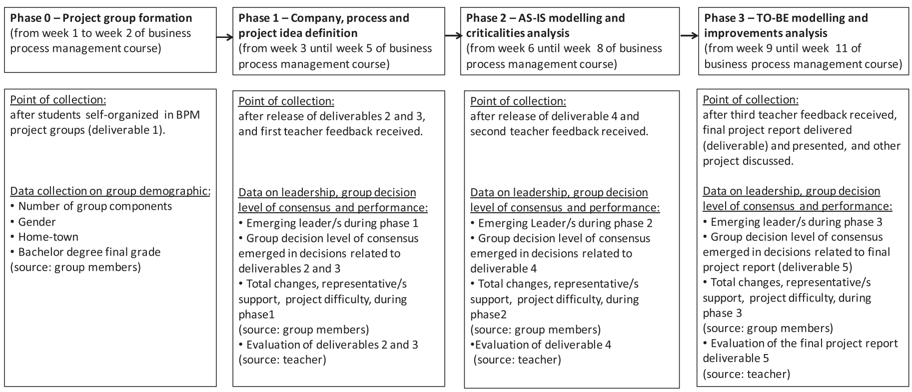
Phase 0 is a preliminary phase in which the groups are set-up. It ends with the submission of the first deliverable (D1), which only contains information about the group and the members names, their contacts, a list of companies suitable for the project, and a provisional team meeting schedule.

Phase 1 requires groups to make important decisions regarding the project (e.g., selecting the company and the process to be analyzed and reengineered, defining the main reengineering idea and a tentative plan of attack) and to accomplish some tasks (e.g., meeting the company managers, interviewing them, analyzing the company's value chain and supply chain). The results of this phase are reported in two separate deliverables (D2 and D3).

Phase 2 requires the groups to identify the list of main criticalities of the company's business process and the list of potential re-engineering solutions to overcome these criticalities. During this phase, the groups also develop a complete map of the business process (AS-IS) and produce deliverable D4 reporting the activities which have been performed in this phase.

Phase 3 requires groups to make the last set of project decisions such as what best practice/s they want to adopt for achieving improvements (they are introduced to a number of possible best practices for business process improvement as listed in Appendix 1), what kind of technology they want to put into the process to enable the reengineered solution, and what Key Performance Indicators (KPI) they want to use in order to assess and demonstrate the potential benefits of the new reengineered business process. Also, the groups perform technical and

Figure 2: Data-collection over the three-month period of the experimental field setting.



time-consuming activities such as designing the new business process map (TO-BE) and conduct simulation and statistical analyses using specific business process simulation and statistics software to measure, test and evaluate the new solution before its actual implementation.

Finally, groups perform some writing, presentation, and discussion activities: they write the last deliverable D5 (which is the final report), they make a presentation of their project in front of the whole class, and they act as discussants of other group BPR projects.

We collected data through four rounds of surveys submitted to the students. The timing for data collection is reported in Figure 2 and simply followed the four phases of the project to capture critical time-points of the project life cycle. In particular, at the end of Phase 0, we collected group members' general data: gender, home-town, and bachelor's degree final grade. At the end of phases 1, 2, and 3 we collected data from the students and from the teacher. To this purpose, the teacher filled a grading rubric for evaluating group decisions that were undertaken by the group up to that moment (see Appendix 2); at the same time, group members were asked to fill out an on-line questionnaire containing questions about our main constructs, i.e., GDC and leadership (see Appendix 3). The multiple time points allowed us to mitigate the issue that perceptional data concerning individual and group dynamics tend to be extremely biased when gathered retrospectively and once only (Huber & Power, 1985).

Given that the purpose of our research is to explore group decision-making behavior, in Table 1 we summarize the two major decisions each group has to make in each phase, while in Table 2 we report an example of the specific decisions made by two groups randomly extracted from our sample.

Measures

Appendix 2 and 3 report the questionnaire items for the dependent and independent variables.

Table 1: Group's decisions to be undertaken in each project phase.

Project Phase	Decision 1	Decision 2
Phase 0 – Group set-up	No decisions in this phase	No decisions in this phase
Phase 1 – Company, business process and project idea	Select the company where the group will conduct the business process re-engineering project	Select the business process the group will analyze and reengineer
Phase 2 – AS-IS mapping and analysis	Identify the business process weaknesses and critical points on which the group will focus for improvement actions	Identify the actions the group will carry out for improvement
Phase 3 – TO-BE design and improvements	Determine the best practices and technologies the group will use for implementing the improvement actions	Determine the key performance indicators that should be measured and showed in the final report to demonstrate the process improvements from the AS-IS to the TO-BE

Dependent variable

Group project performance. Our hypotheses link group decision behavior to group project performance. Accordingly, group project performance, that is our dependent variable, reflects the effectiveness of group decisions. The measurement of group project performance was based on project scores assigned by the teacher, which followed the Italian universities' grading system ranging from 18 up to 30 cum laude. The teacher assesses and assigns grades at the end of each of the three main project phases (phases 1, 2, and 3). In all three assessments, the teacher evaluated the quality and the level of evidence-based nature of the project releases related to the group decisions undertaken by the time of assessment (see Appendix 2). We finally calculated group project performance in each phase as the average of all scores assigned by the teacher to the project-group performance dimensions evaluated in that phase. To limit a potential bias in group performance measurement introduced by possible confounding between teaching the course and conducting the research, we completely separated the data collection activities related to dependent and independent variables. While the teacher assessed the performance measure, all the remaining data related to the other variables of the conceptual model were collected by an independent researcher, who directly conducted the three-rounds of collecting survey data, elaborated the related measures, and built the final dataset at the end of the project after teacher evaluations.

Independent variables

Group decision level of consensus (GDC). Studies on decision consensus capture the level of team members' agreement/disagreement about a specific issue and

Table 2: Business process re-engineering projects examples of group decisions undertaken in each project phase.

Decision 1 (Phase 1)				Decision 2 (Phase 1)	Decision 1 (Phase 2)	Decision 2 (Phase 2)	Decision 1 (Phase 3)	Decision 2 (Phase 3)
Team	Industry Sector	Company Main Product/Service	Firm Size (Number of employees)	Business Process	Business Process Criticalities	Improvement Actions	Best Practices	KPI
Team 1	Wholesale trade	Wholesaler of grocery products (food, beverage and house-hold products)	11	Replenishment order management	High level of: (i) Out of Stock (OOS) (ii) Average Order Cost (iii) Average Order Time Main cause individualized: (i) Replenishment order based on worker experience rather than on an optimal reorder policy	(i) Alert cashers to pay attention to register the correct inventory code of sold products in order to reduce error of outbound products registration (ii) Ask suppliers for electronic receipt of delivered products in order to reduce time and error of inbound products registration (iii) Adopting a software that implements a fixed order point policy	(i) Control addition (ii) Integration with Suppliers (iii) Task automation (iv) Integral technology	(i) Average cost of order launch (ii) Holding Inventory cost (iii) Out of Stock cost (iv) Service level (v) On Shelf Availability
Team 2	Manufacturer	Manufacturer of painting products for the construction industry	41	Customer order management	High level of: (i) Average flow time of catalog-based products (ii) Average flow time of customized products (iii) Number of faulty processed orders due to communication mistakes	(i) Appointing one worker of the laboratory unit as responsible for the handling of customized orders in such a way that he/she directly communicates with the customer without the intermediation of the sales department manager (the idea is to reduce the likelihood of communication mistakes),	(i) Case manager (ii) Buffering (iii) Integral technology	(i) Average flow time of catalog-based products (ii) Average flow time of customized products (iii) Number of faulty

(Continued)

Table 2: Continued

Decision 1 (Phase 1)				Decision 2 (Phase 1)	Decision 1 (Phase 2)	Decision 2 (Phase 2)	Decision 1 (Phase 3)	Decision 2 (Phase 3)
Team	Industry Sector	Company Main Product/Service	Firm Size (Number of employees)	Business Process	Business Process Criticalities	Improvement Actions	Best Practices	KPI
					(iv) Number of faulty processed orders due to transcription mistakes	who will have more time to dedicate to catalog-based orders (the idea is to reduce the transcription mistakes). (ii) Information sharing between the production, the inventory, and the sales department by implementing a software where both orders to be produced and orders produced can be visualized by the three units, eliminating the need for paper or electronic message exchange (the idea is to reduce the order flow-time and transcription mistakes)		processed orders due to communication mistakes (iv) Number of faulty processed orders due to transcription mistakes

measure it using Likert-scale responses to several items related to the specific issue, and then calculate over those responses either a group coefficient alpha (Jehn & Mannix, 2001) or an aggregated standard deviation over the standard deviations related to the team members' responses (Knight et al., 1999; Mohammed & Ringseis, 2001). However, because we want to capture also consensus reaching mechanism behaviors (related to the GT and AP phenomena), we developed a different measure for GDC. Specifically, as described in the literature review section, GT-like consensus behavior occurs when group members striving for unanimity (i.e., low divergence between the individual and the group thinking) override their motivation while not realistically appraise alternatives to a unanimous decision and are finally satisfied to experience a sense of unanimity rather than looking for a really motivated and discussed decision. On the other hand, AP-like consensus behavior occurs when group members, and hence groups as a whole, take decisions in contradiction to what they really want to do (i.e., high divergence between the individual and the group thinking); at the end, "no one is satisfied about the trip to Abilene!" It means that both GT- and AP-like consensus behaviors manifest through a certain level of group members' satisfaction about the group decision (high in GT and low in AP), given a certain level of divergence between the individual private position with respect to that decision and the decision itself (low in GT and low in AP).

Consistent with this conceptualization, we developed an indicator that measures the ratio of group members' satisfaction over the divergence between individual and group decision. In fact, the higher this ratio is, the higher the likelihood of GT-like consensus behavior (i.e., high satisfaction, low divergence between individual and group decisions); the lower it is, the higher the likelihood of AP-like consensus behavior (i.e., low satisfaction, high divergence between individual and group decisions). In order to measure GDC, we selected two main critical decisions the group made during each phase of the project (1, 2, and 3), as reported in Table 1. For each of these decisions we asked each group member (i) if the final group decision coincides with what she/he would have personally chosen (i.e., the group member divergence from the final group decision); (ii) how much she/he feels happy about that final decision (i.e., the group member satisfaction about the final group decision). We considered response to question (i) as a dummy variable that takes the value 1 if the respondent answered "yes"; 0 otherwise. Regarding response to question (ii), its measure was based on a 5-point Likert scale ranging from 1, "Very Dissatisfied," to 5, "Very Satisfied." We finally calculated the GDC in each phase as the ratio of the average of group members' perceived satisfaction over the average of group members' divergence from the final group decision.

Shared leadership. Shared leadership captures the distribution of leadership influence across multiple team members. We used the roster method to collect data on shared leadership within each BPR group (Wellman & Berkowitz, 1988). In each team, respondents were asked to indicate the members they perceived to be leader/s. Respondents were free to nominate as many or as few leaders as they deemed appropriate. This operationalization is consistent with the classic sociometric work on leadership in teams (e.g., Stogdill, 1948; Shaw, 1964), and it is also consistent with the theoretical conception of leadership as a phenomenological construct: a leader is someone who is perceived as such by others (Calder, 1977; Pfeffer, 1977; Meindl, 1993). The questionnaire does not specify what we meant by the term "leader" because we are interested in capturing respondents' naive theories of leadership (Lord & Maher, 2002). Moreover, because according to Carson et al. (2007), p. 1220) "teams with high levels of shared leadership may also shift and/or rotate leadership over time, in such a way that different members provide leadership at different points in the team's life cycle and development," in the last two rounds of surveys, we alerted group members that it was not mandatory to report the same member/s as marked in the previous surveys, unless they perceived that the leader/s was the same (see Appendix 3). Finally, following Neubert (1999), we computed shared leadership by dividing the number of members who had received at least one leadership nomination within the team by the total number of members. A high value of shared leadership indicates a more distributed leadership across team members.

Control variables

In order to control for other factors influencing group project performance, we introduced in our model several additional variables. To account for differences in the level of difficulties faced by a group, both objective and subjective, we included three variables: total changes, company support, and project difficulty. The first two variables model objective sources of difficulties, while perceived project difficulty models their subjective component. *Total changes* refer to the times the group changed the chosen company and/or process during the BPR project. Indeed, past experiences demonstrate that it can happen that during the development of the project the group realizes that a different but more appropriate process should be selected, and/or that unfortunately the company abandons the project and a new company needs to be contacted. Because it is an objective and quantitative measure, to avoid respondent-bias and/or random error we checked that the within-group standard deviation of responses on company's changes as well as on process changes, was not significantly different to zero in all groups and at each time of data collection. Because this criterion was satisfied, we assume validity of this measure. *Company support* measures the availability and help of a company's contact person in collaborating with the project group. We measure this variable through a two-item 5-point Likert scale, asking for both the level of availability and the level of concrete support of a contact person in the project company (both are based on the perception of team members). *Perceived project difficulty* represents a subjective measure of how much the group members comprehend the BPR project as difficult. It can depend on several reasons, such as perception of overloading and consequent stress, as well as on members' personal traits (e.g., their level of anxiety). We measure this variable by asking respondents to indicate how much they find the project difficult using a 5-point Likert scale.

Given that data were collected in three time points, we also control for the *time of observation* effects by indicating time at the end of phases 1, 2, and 3 as 1, 2, and 3, respectively. We control for time on performance because the groups tend to perform better in the last phases of the project because they know that their final grade mostly depends on the final report delivered to the teacher. Also, we control for group *gender-homogeneity*, calculated as the ratio between the number of members with the most frequent gender in the group and the total number of group members; for *home-town homogeneity*, calculated as the ratio between the number of members coming from the most frequent home-town in the group and the total number of group members; and for group *bachelor degree final grade* calculated as the average of the bachelor degree score of all members constituting the group (Williams & O'Reilly, 1998).

ANALYSES AND RESULTS

Aggregation of Individual-Level Variables at Group-Level

To derive some of the group-level variables included in the conceptual model, we aggregated individual-level observations. We initially had 377 individual-level observations (collected from 141 students), which were grouped into 100 group-level observations (related to 34 groups). We adopted two different approaches to aggregation depending on what we were interested to capture at the group-level.

We adopted an additive aggregation composition model (Chan, 1998) for those variables of which we were interested at measuring the degree to which they are exhibited, in aggregate, within the group, regardless whether group members agreed about them. For example, for computing the construct GDC within a group, we need to know (calculate) the average level of satisfaction among group members. This is irrespective of how much each member agrees with this level. We used this approach for the variables: group members' satisfaction on group decisions, divergence between individual and group decisions, and perceived project difficulty. Accordingly, for such measures within-group agreement was not a statistical prerequisite for aggregation. These measures conform to what Klein and Kozlowski (2000) refer to as configural constructs. Specifically, as they argue, "constructs of this type capture the configuration of individuals' characteristics within a unit. Unlike shared unit properties, however, configural unit properties are not assumed to coalesce and converge among the members of a unit" (p. 30). Concretely, we aggregated satisfaction, divergence, and perceived project difficulty to group-level by computing the average of the values among members within the group. Of the 100 group-level observations derived from aggregating the collected 377 individual-level observations, we eliminated 1 group-level observation because it resulted from the response of only one of its members. Moreover, we eliminated 14 group-level observations where no divergence between the personal and the group view about decisions undertaken by the time of collection was declared. The rationale under the first elimination is the incoherence of basing a group-level observation on the response of just one of its group-members. The second elimination, instead, avoids the risk to consider in the analysis some groups in which the process of reaching the consensus could not be observed because there was no gap at all between initial individual thoughts and final group decision. Thus, our final sample consists of 85 observations related to 33 BPR groups.

On the other side, we adopted a direct consensus model for the items reflecting *Company support* (i.e., company availability and help), because we wanted to compute an objective measure for this construct as much as possible (Chan, 1998). Because in this kind of aggregation model, demonstration of sufficient within-group agreement is a necessary precondition for the aggregation of micro-level (e.g., individual-level) measures to represent macro-level (e.g., group-level) constructs (Cohen, Doveh, & Eick, 2001), we checked for *inter-rater agreement (IRA)*, that references the interchangeability among raters and addresses the extent to which raters make essentially the same ratings (Kozlowski & Hattrup, 1992). We computed three measures to assess IRA, such as those discussed by Boyer and Verma (2000): the *percentage method*; the *ratio method*; the *interclass-correlation coefficient (ICC) method*. While the ICC method has the advantage of testing for statistical significance (it checks if within group variance is negligible compared to between group variance), it is not always a reliable measure of whether people within organizations agree or not because it basically measures interorganizational heterogeneity (Ketokivi & Schroeder, 2004). For this reason, we also adopted the other two methods. From the analyses (not reported for sake of brevity), we conclude that the IRA in the data is good for the *Company's contact person availability* item but not for the *Company's contact person concrete support* item. Thus, we decided to aggregate at group level just the first one as measure of *Company support*, by computing the average among the group members' responses. Because *Company support* together with *Total changes* and *Perceived project difficulty* are potential expressions of the level of difficulty faced by the group, we checked for the significance of pairwise correlations among these variables, but we found these not to be statistically significant, so we did not aggregate them.

Endogeneity Concerns

We hypothesize (H1) that there exists a causal relationship between *Group Decision Level of Consensus* (GDC) and *Group Project Performance*, and in particular that the variance of GDC drives the variance of performance through an inverse U-shaped relationship. A concern when evaluating such a causal relationship is the potentiality for a "simultaneity bias" as a potential source of endogeneity. As reported in Roberts and Whited (2013, p.11), "simultaneity bias occurs when y and one or more of the x's are determined in equilibrium so that it can plausibly be argued either that x_k causes y or that y causes x_k." In particular, *Group Project Performance* may also influence the level of GDC, in such a way that the higher the performance the higher the group members' feeling of satisfaction about decisions and accordingly the higher the GDC. Indeed, as reported by Bendoly et al. (2010a), high level of group decision quality and operational performance could indirectly increase the likelihood of GT occurrence (i.e., high level of GDC) and either increase or decrease the likelihood of AP (i.e., low level of GDC), through the mediating effect of situational pressure, group blame, and cohesion. As an example, better performing groups (i.e., high level of decision quality and operational performance) are likely to increase the level of "fake" consensus about group decisions because the acknowledgment of having performed well in the past could increase their feeling of satisfaction about group decision, thus making our estimates of the impact of GDC on *Group Project Performance* biased.

To address this endogeneity concern, as Freedman (1991, p. 292) notes, statistical technique is rarely a substitute for good empirical design; so, in order to control for this potential informant bias we *ex ante* designed the data collection in such a way that the students were not aware about the project grades assigned by the teacher in the three project phases, and the teacher, that is also one of the researcher involved in this study, was not aware of the students responses, up to the end of data collection. By this, we avoid that the variance of GDC would be driven by the variance of performance, either because the acknowledgement of high/low group project performance would induce its members to over/underestimate their perceived satisfaction, or because the acknowledgement of high/low GDC would induce the teacher/researcher to underestimate the performance. Accordingly, such a "double-blind" data-collection of the independent and dependent variables, at least theoretically, allowed us to avoid informant bias that would lead to "simultaneity bias" as a source of endogeneity. Moreover, the use of two different sources of information for the dependent and the independent variables by using the group members as source of the independent variable and the teacher/researcher as source of the dependent, reduced the likelihood of correlated errors between the independent and dependent variables. As suggested by Ketokivi and Schroeder (2004, pp. 254), "correlated errors are likely in the case in single-informant surveys, where the same informant is the source of both the independent and dependent variables."

However, GDC is still likely endogenous in our conceptual model because it can be influenced by shared leadership. Given this endogeneity, using the raw values of GDC in the regression may lead to confounds of the effects of GDC with the moderator (i.e., shared leadership). As in Bai, Sheng, and Li (2016), we use two-stage least square regressions to correct for endogeneity, given that the supposed endogenous independent variable (i.e., GDC) and the dependent (i.e., performance) are continuous (Hamilton & Nickerson, 2003).

In Stage 1, we regressed GDC on *Shared Leadership*. The regression results showed that GDC related positively to *Shared Leadership* (β = -13.29, $p < .01$). These results confirmed the significant impacts of *Shared Leadership* on GDC in group-decision making processes and the need to use two-stage regression models to address potential endogeneity among the predictors. In Stage 2, we used the residuals (observed value minus predicted value) as indicators of GDC. These residuals represent the portion of GDC that remained unexplained by *Shared Leadership* and we use them as instrument of GDC in the second stage regression. The practice of using residuals as an instrument in the second-stage regression is quite common (e.g., Zhou & Li, 2012; Bai et al., 2016) as it allows correcting for a potential endogeneity issue because such residuals are free of the influence of the predictors used in the first-stage regression. We then used the residuals to construct the interaction terms and entered them into the second-stage regression models.

Model Estimation Results

Table 3 provides the summary statistics on the key variables used in the final analysis. Pairwise correlations are computed using Pearson calculations, whose level of significance are determined using a one-sided *t*-test. The correlation between *Shared Leadership* and GDC is negative (r = -.358, $p < .01$) and its correlation with *Group Project Performance* is positive (r = .247; $p < .05$).

Table 3: Descriptive statistics and pairwise correlations (N = 85 from 33 business process re-engineering groups).

		Mean	*SD*	1	2	3	4	5	6	7	8	9	10
1	Group project performance	26.35	2.12	1									
2	Group decision level of consensus	10.40	8.77	0.00	1								
3	Shared leadership	0.77	0.24	0.25**	−0.36***	1							
4	Company support	3.80	0.98	0.27**	−0.02	0.13	1						
5	Total changes	0.28	0.59	−0.31***	−0.08	0.00	−0.17	1					
6	Perceived project difficulty	3.65	0.52	0.13	0.07	−0.18	−0.05	0.12	1				
7	Time of observation	1.92	0.82	0.34***	−0.14	0.04	−0.11	−0.25**	0.29**	1			
8	Gender-homogeneity	0.75	0.20	−0.13	0.12	−0.06	−0.21*	−0.11	−0.23**	0.00	1		
9	Home-town homogeneity	0.38	0.30	0.04	−0.06	0.08	0.05	−0.23**	−0.22**	0.08	−0.04	1	
10	Bachelor degree final grade	97.24	5.75	0.23**	0.03	0.01	0.06	−0.04	0.01	0.08	−0.02	−0.10	1

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$.

The longitudinal investigation of our research led our data to be hierarchically nested (repeated measures nested within groups). As observed by Hofmann (1997), the nested nature of longitudinal data consists of multiple observations within a unit (i.e., the observations related to each BPR group at time 1, 2, and 3) and a sample of multiple units (i.e., the observations related to several BPR groups). As data were collected three times from the same sources (BPR groups), the observations of our measures are clustered into BPR groups. Due to the nested nature of our data, we use a hierarchical linear model (HLM) to test our hypotheses (Raudenbush, Bryk, & Congdon, 2004). In order to assess the appropriateness of HLM, we follow the suggestion by Hofmann et al. (1997) and tested for significant between-group variance before analyzing the full model. We conducted an ICC analysis and associated chi-square tests to check on variance in our measure of Group Project Performance between and within groups. We first ran the null model that had no predictors. The null model is analogous to an analysis of variance in that it partitions variance in Group Project Performance into within- and between-group components. The estimate of the ICC is computed as the proportion of between-groups variance over the total variance. The result (ICC = 0.36; $p_{\chi 2} < .01$) indicates that 36% of the variance in Group Project Performance significantly resides between groups, and 64% within groups, hence suggesting the need to account for time effects through HLM. Accordingly, we used the mixed model linear regression in STATA to test our hypotheses.

Given that our model involves both additive and multiplicative effects, we present the regression results in a gradual manner to illustrate the incremental predictive power associated with the addition of the main effects and the interaction effects in the analysis (see Table 4). Model 1 represents the baseline model with control variables only, Model 2 introduces the independent variable (i.e., GDC) of first and second order, Model 3 introduces the direct effect of the moderator variable, and Models 4 introduces the interaction terms between GDC and shared leadership of first and second order. To mitigate multicollinearity concerns, we centered the independent variables before creating higher order and interaction terms. Following a backward elimination stepwise regression approach (Mao, 2004), we eliminated from our regression models additional controls not found to be statistically significant in our analysis. We therefore only include controls that revealed to be significant, along with our main independent variables. Accordingly, the following controls have been excluded: *perceived project difficulty, gender-homogeneity, home-town homogeneity, bachelor degree final grade.*

H1 posits that group project performance has an inverse U-shaped relationship with group decision level of consensus. To test this hypothesis, we examine both the linear and the quadratic effects of GDC in Model 2, thus we use a parabolic function (quadratic polynomial function) to model the U-shaped relationship. We also tried other polynomial functions (see the next section on robustness check) but the quadratic model showed the best fit. The regression results indicate a significant positive effect for the linear term of GDC ($\beta = 0.1, p < .01$) and a significant negative effect for the squared term of GDC ($\beta =$ -0.005, $p < .01$) on group project performance. To get a deeper understanding of these effects, we plot the fitted values for the relationship between GDC and group project performance. As

Table 4: Analysis results (N = 85 from 33 BPR groups).

	HLM Regression			
	Group Project Performance			
	Model 1	Model 2	Model 3	Model 4
Independent variables				
Group decision level of consensus (GDC)		0.1 (0.03)***	0.09 (0.03)***	0.07 (0.05)
GDC^2		−0.005 (0.00)***	−0.004 (0.00)**	−0.004 (0.00)**
Shared leadership (SL)			1.66 (0.94)*	0.99 (0.98)
GDC X SL				−0.2 (0.19)
GDC^2 X SL				0.02 (0.01)**
Control variables				
Total changes	−0.77 (0.36)**	−0.75 (0.34)**	−0.77 (0.33)**	−0.65 (0.33)*
Company support	0.49 (0.23)**	0.5 (0.23)**	0.49 (0.22)**	0.49 (0.22)**
Time of observation	0.82 (0.21)***	0.92 (0.20)***	0.90 (0.20)***	0.89 (0.19)***
Constant	23.11 (1.12)***	23.11 (1.08)***	21.91 (1.24)***	22.58 (1.27)***
Log likelihood	−167.81	−163.4	−161.91	−159.95
Wald chi square	29.80***	43.65***	47.54***	54.39***
R^2	0.29	0.41	0.41	0.46

R^2 estimates based on Kreft & Leeuw (1998), and Singer (1998).
$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$, *SE* in parenthesis.

Figure 3: Inverse U-shape relationship between group decision level of consensus (GDC) and group project performance.

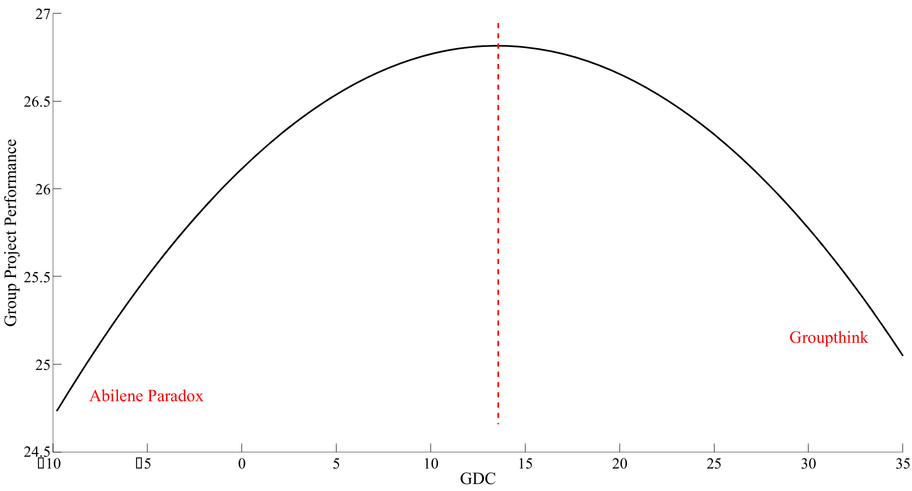
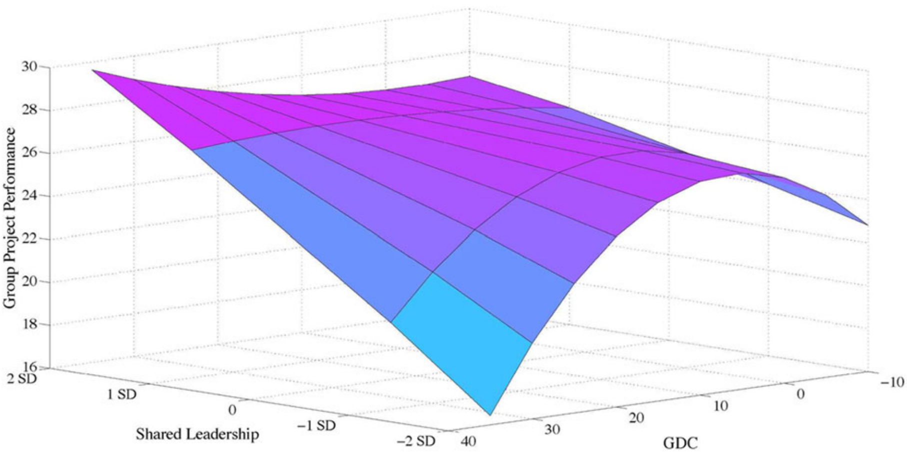


Figure 3 indicates, group project performance improves with GDC up to a certain point beyond which an increase in GDC is associated with a decrease in group project performance. To decrease multicollinearity between the interaction term and its corresponding main effects, we mean-centered the GDC variable; this explains its range from negative to positive in Figure 3. Results provide thus support for H1 and suggest that low values of GDC are associated with low performance: on the left side of the parabola we have groups with AP-like consensus (low satisfaction, high divergence, low performance); also, high values of GDC are associated with low performance: on the right side of the parabola we find groups with GT-like consensus (high satisfaction, low divergence, low performance).

H2 posits that shared leadership moderates the relationship between GDC and group project performance. In fact, we are expecting that shared leadership strongly reduces or even eliminates the occurrence of GT and AP phenomena. To confirm H2, we interpret the results from Model 4, consistent with other studies that hypothesize moderation effects on nonlinear relationships (e.g., Lapré & Tsikriktsis, 2006; Mishra, Chandrasekaran, & MacCormack, 2015). The interaction of shared leadership with the linear term of GDC ($\beta =$ -0.20, $p >.1$) is negative and not significant, whereas that with the quadratic term (β = 0.02, $p <.05$) is positive and significant. To better understand this interaction effect, we follow Aiken and West (1991) and plot the relationship between GDC and performance at different levels of shared leadership (-2 *SD*, -1 *SD*, $+1$ *SD*, $+2$ *SD*). The surface plot in Figure 4 indicates that, at lower levels of shared leadership, GDC has an inverse U-shaped relationship with group project performance. As shared leadership increases, the curve progressively flattens and then reverses in shape. Similar results for moderation effects in quadratic relationships have been found in previous studies (e.g., Mihalache, Jansen, Van Den Bosch, & Volberda, 2012;

Figure 4: Interaction of shared leadership and group decision level of consensus (GDC) on group project performance.



Mishra et al., 2015), where the change in the shape of the curve has been associated with a strong effect of the moderator variable. H2 is therefore supported. Finally, the control variable *total changes* has a significant and negative effect on group project performance, while *time of observation* and *company support* a positive effect, as presumed.

Robustness Checks

To verify the robustness of our results, we performed some additional tests. First, to provide evidence that a U-shaped relationship between GDC and *Group Project Performance* fits data better than other specifications, we tested for a potential cubic relationship, but results were not statistically significant.

Second, as in Mihalache et al. (2012), we followed the procedure advanced by Aiken and West (1991) and conducted a simple slope analysis to test the statistical significance of the relationship between GDC and performance at different levels of shared leadership (-2 *SD*, -1 *SD*, $+1$ *SD*, $+2$ *SD*). Results indicate that the simple U-shaped slope of the regression curve increases (from negative to almost null values) when *Shared leadership* increases. It is negative and significant at very low ($\beta =$ - 0.012, p <.05) and low levels of shared leadership ($\beta =$ - 0.008, p <.01). Conversely, the simple slopes are not significant (p >.1) for very high and high levels of shared leadership. These findings provide additional evidence in support of both the U-shaped specification and that the inverted U-shaped slope would almost disappear (flatten) in groups with high levels of shared leadership.

Third, as in Mishra et al. (2015), we used the two-steps residual centering technique recommended by Lance (1988) to reduce multicollinearity concerns associated with including higher order interaction terms. In the first step, we got the residual variables obtained by regressing each interaction term against its constituent main effect terms. In the second stage, we replaced the interaction terms in the original model with the residual variables resulting from the first step. Such a replacement eliminates the shared covariance between the main effects and the interaction terms, thereby reducing multicollinearity concerns. This robustness check confirmed the results of our main analyses. Finally, we reconducted the HLM analysis three times using three randomly selected subsamples (90%, 80%, and 70% of the original sample) and we found the same empirical findings as when using the full sample.

DISCUSSION AND CONCLUSION

Summary of Findings

This study focuses on group decision making in BPR projects. In particular, we examine the relationship between the level of consensus reached in operational decisions and project performance; we also investigate whether and how this linkage changes in groups where the leadership is shared among several members rather than being concentrated on one or few members. We conceptualize the level of GDC as the level of group members' satisfaction about a group's decisions, for a given level of divergence between what the members would individually decide and what their group decides. Grounding our argumentations on group decision-making literature, we show that this way of abstracting the group consensus allows us capturing GT and AP behavior in group decision-making. Then, we develop and test hypotheses that examine both the direct impact of GDC on group project performance and the role played in this regard by the level of shared leadership inside the groups.

Using primary data stemming from a field experiment, we find support for a curvilinear relationship between project performance and the level of GDC. Furthermore, we find that in groups characterized by higher levels of shared leadership this relationship is moderated in such a way that the curve appears reversed, signaling that both GT-like and AP-like negative group decision-making consensus behaviors are mitigated. That is, at higher levels of shared leadership, group project performance benefits associated to deeper group members' interactions and discussions (Roch, 2007) allow to eliminate the occurrence of negative group phenomena in group decision-making processes (i.e., GT and AP). Accordingly, a higher level of GDC (high satisfaction for low divergence) will not be a signal of GT-like consensus behavior but represents the proof of a discussed and really shared decision. Similarly, a lower level of GDC (low satisfaction for high divergence) will not be a signal of AP-like consensus behavior but represents the fact that, even if the final group decision was not really shared (a real consensus on the decision has not be achieved), members did not conform to the thinking of an inconsistent leader (who drives the group to Abilene) but the discussion was really animated and the decision was examined in depth by members with committed divergent opinions.

Contributions to Theory

We believe our study provides theoretical contributions to the interdisciplinary group decision-making literature and to the management literature asking for additional effort in reducing the lack of behavioral studies in operational contexts (Bendoly & Hur, 2007; Croson, Schultz, Siemsen, & Yeo, 2013). In fact, management research is limited in the area of operational group behavior and dynamics, which is indeed where this article places its focus.

First, we introduce a new concept, and related measure, of consensus in group decision-making (GDC) as the level of group members' satisfaction with a specific decision for a given divergence between this decision and the group members' private views on that. We provide an argumentation demonstrating that this way of abstracting the level of consensus reached allows considering the possible occurrence of GT- and AP-like consensus behavior when making decisions in operational teams. We accordingly extend the recent group conflict management literature by showing that not just conflict, but even extreme consensus has potential drawback in group decision-making process, and thus future studies have to consider these double aspects when investigating CRP for effective decision.

Second, our theoretical argumentation and empirical results add new knowledge about the consequences of GDC on BPR project performance. More specifically, we provide support for the AP and GT theories (Harvey et al., 2004) and demonstrate that both AP-like and GT-like consensus have negative implications for group performance not just in top management teams but also in operational contexts. This is the first study that considers both group decision conflict (AP-like consensus behavior) and unanimity (GT-like consensus behavior) issues in operational teams.

The third contribution of our study comes from the examination of the moderating effect of shared leadership on the relationship between GDC and group project performance. While most studies in the field of operations management, information system management, and organization science tend to focus on benefits provided by vertical rather than shared leadership (e.g., McFadden et al., 2009), our study emphasizes and empirically shows the benefit of the dispersion of leadership over multiple internal group members. Also, more conceptually speaking, despite the key role played by leadership in determining GT and/or AP (i.e., being an antecedent), no organizational behavior studies exist focusing on the role played by shared leadership in contrasting them (i.e., being a moderator). Our study, contributes to this literature by providing empirical evidences of the moderating effect played by shared leaderships in reducing the effect of the two negative phenomena.

This finding also contributes to clarifying the discussed debate in the concurrence-seeking literature regarding the negative or positive impact of group concurrence seeking behavior (GT-like consensus behavior) on the effectiveness of group decisions (Park, 2000; Chen et al., 2009). Our findings suggest that the effectiveness of this behavior is strictly connected to the level of leadership dispersion inside the group, implying that a high level of shared leadership implements virtuous circles of team effectiveness through the benefits of social networking and learning, dampening the risk of conformity and misperception of the collective reality, and fostering discussion and in-depth examination of every decision. This result complements recent findings that underline the important role played by factors such as information sharing and psychological safety (Bendoly, 2014; Tsai & Bendersky, 2015), collectivistic behavior (O'Neill & McLarnon, 2017), and mutual evaluation (Dong et al., 2016). While, in fact, those studies implicitly assume these factors as antecedents of positive group work performance, we contribute by explaining the reason that underlies such a relationship: as a consequence of shared leadership within the group, these factors act as moderators between GT-like consensus behavior and the effectiveness of the decision; in other terms, shared leadership explicates the positive side of group concurrence seeking behavior.

Contributions to Practice

This study provides managers with several insights for an effective management of temporary operative groups, avoiding pitfalls dangerous for group project performance.

Specifically, the inverse U-shaped relationship between group project performance and GDC suggests external group managers to participate in group discussions of key project issues to check for dangerous concurrence-seeking behaviors. These can manifest in two cases. In one case (AP-like consensus behavior), it happens that in group discussions too few members express their views before they concur with the final decision, and show several nonverbal cues suggesting anger (e.g., frowns, tense movements; Manz & Sims, 1982) after the group decision has been made. In the other case (GT-like consensus behavior), group members during discussion, avoid any form of critical thinking while just providing support for views expressed by others and omit alternative solutions, and finally show high level of satisfaction. Accordingly, besides paying attention to the several symptoms which may arise during the group decision-making process (the literature on AP and GT has already identified and classified these, e.g., Kim, 2001; Harvey et al., 2004), we suggest managers to check for two further signs: (i) the intensity of group members' participation and expression during group decision-making discussions, and (ii) the level of group members' satisfaction, as expressed by nonverbal cues. Too much or too little of these two elements may be an alerting signal for an ineffective group decision-making process.

Besides providing suggestions on how to detect such negative group decision making behavior, this study allows us to offer insights on how to reduce the propensity of operational groups in incurring in GT and AP negative behaviors and thus how to increase the likelihood of project success. Our results, indeed, indisputably show the beneficial effect of shared leadership in moderating the negative implications of very low or very high levels of GDC on project performance. When shared leadership is high, even extreme (thus suspicious) levels of consensus are symptoms of fruitful discussion and efficient decision-making processes, meaning the shared leadership is able to mitigate GT and AP effects. This clearly means that managers are asked to invest in training and coaching programs aimed at increasing leadership skills of all team members involved in operational projects. In other words, the members of an efficient BPR project team should not just possess hard and technical skills (e.g., business process mapping skills, production flow analysis skills) but they all should also possess soft skills usually requested for rather managerial positions (e.g., social graces, communication, language, interpersonal skills, driving, motivating and supporting each other, leadership). By this way virtuous circles of social networking and social learning can be implemented, allowing the group to develop in favor of the project outcomes. Discussions about important project issues would benefit from sharing critical ideas and making a really joint and well-discussed group decision. An example is provided by the success of self-managing teams in flat organizational structures (Manz & Sims, 1987) that demonstrate the need for leadership originating from within a team as opposed to that originating from a single individual elevated by hierarchy (Carson et al., 2007). Self-managing teams are teams where leadership responsibilities are shifted from formal managers to team members. These members are employees who have high levels of expertise and seek autonomy in how they apply their knowledge or skills (DeNisi, Hitt, & Jackson, 2003). They only need to be trained more on leadership skills to increase the level of shared leadership within the group.

Limitations and Future Research

Our study has several limitations that may also be a source for future research.

First, our study relies on only one measure of shared leadership. Although our measure is consistent with other sociometric works on leadership in teams (Stogdill, 1948; Shaw, 1964) and with the theoretical conception of leadership as a phenomenological construct (Calder, 1977; Pfeffer, 1977; Meindl, 1993), several other measures of shared leadership exist (e.g., future research can take into account the interrelated network emerging among team members through a network analysis) and could be used to provide supporting or complementary views to our findings.

Second, our conceptual model focuses on AP and GT phenomena as antecedents of conflict/consensus processes and poor decision making. We do not consider other potentially harmful and somehow overlapping phenomena such as majority influence and conformance pressure (Zhang, Lowry, Zhou, & Fu, 2007), cultural conflict (Maznevski, 1994), and voice behavior (Morrison, 2011) that have been demonstrated to create poor group decisions. Also, we look at shared leadership as a mechanism to contrast AP and GT, but we disregard other mechanisms, such as collaborative technology (Lowry, Roberts, Dean, & Marakas, 2007), coordination technology (Ren, Kiesler, & Fussell, 2008) temporal coordination (Montoya-Weiss, Massey, & Song, 2001), and knowledge sharing (Sillito Walker & Bonner, 2018) as dampers of these phenomena. Future studies should consider how these other undesirable phenomena overlap and, eventually, influence each other and how to design counterbalancing mechanisms, such as shared leadership, coordination, and collaborative technology.

Third, we consider temporary operational groups. In fact, our units of analysis are provisional teams working in a BPR project, whose life cycle corresponds to that of the project. These groups experience a high level of stress, mainly driven by restrictive deadlines and constraints (e.g., budget). We, indeed, put in place a provocative situational context where social factors become critical and the occurrence of negative and paradoxical behavior in group-project decision-making processes is likely. However, it would be necessary for future research to explore possible differences between temporary versus stable operational teams regarding the consequences of reaching consensus in group decision-making.

Fourth, even if our data set is longitudinal through different phases of the project and we control for the time the data was collected, a deeper understanding of group decision making dynamics in BPR projects should take more deeply into account the different phases of project life cycle. When and how does a specific behavior manifest in a particular phase rather than in another? Answering this question would surely provide interesting insights for managers who oversee BPR teams.

Fifth, we did not conduct our research with real business teams and our sample was composed by team of students carrying out BPR projects. Although laboratory and field experimental methods are a well-established paradigm in organizational behavior research, we are conscious that real business teams are part of an organization, which is usually characterized by specific and own processes, norms, history, geography, and culture. These organizational aspects should be taken in consideration in future studies that may investigate their impact on the occurrence of GT and AP phenomena in operational teams. For example, the impact on group decision behavior of tacit versus explicit corporate norms, of domestic versus international teams should be definitively considered.

Sixth, even if we control for gender-homogeneity and hometown-homogeneity we do not consider other group composition effects, that could likely influence group performance. For instance, individual characteristics such as abilities, opinions, or personal traits could affect group composition (Moreland, Levine, & Wingert, 2013) and in future studies, issues related to group composition, group structure, and the ecology of groups should be definitively considered when studying group conflict/consensus versus group performance.

Finally, we use the teacher evaluation to measure group project performance. We feel that future research should collect evaluations from multiple respondents from the company where the BPR project is being conducted.

We believe that the insights developed in this study will not only help managers to put in place, to train, and to coach more effectively (in terms of decision-making) operational teams, but also to interpret and predict the quality of the group's decision by observing the way in which operational decisions are made and the level of consensus reached. Also, we believe our findings may motivate researchers who are interested in group dynamics to investigate other factors that moderate the relationship between GDC and group project performance.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix 1
Appendix 2
Appendix 3

REFERENCES

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, London, UK: Sage.

Alby, F., & Zucchermaglio, C. (2006). 'Afterwards we can understand what went wrong, but now let's fix it': How situated work practices shape group decision making. *Organization Studies*, *27*(7), 943–966.

Appan, R., Mellarkod, V., & Browne, G. J. (2005). The role of the Abilene Paradox in group requirements elicitation processes. *AMCIS 2005 Proceedings*. Paper 185. http://aisel.aisnet.org/amcis2005/185.

Asch, E. S. (1955). *Social psychology*. New York: Prentice Hall.

Avolio, B. J., Jung, D. I., Murry, W., & Sivasbramaniam, N. (1996). Building highly developed teams: Focusing on shared leadership process, efficacy, trust, and performance. In M. M. Beyerlein & D. A. Johnson (Eds.), *Advances in interdisciplinary study of work teams: Team leadership*. Greenwich, UK: JAI Press, 172–209.

Bai, X., Sheng, S., & Li, J. J. (2016). Contract governance and buyer–supplier conflict: The moderating role of institutions. *Journal of Operations Management*, *41*, 12–24.

Bamford, D., & Griffin, M. (2008). A case study into operational team-working within a UK hospital. *International Journal of Operations and Production Management*, *28*(3), 215–237.

Bendoly, E. (2014). System dynamics understanding in projects: Information sharing, psychological safety, and performance effects. *Production and Operations Management*, *23*(8), 1352–1369.

Bendoly, E., & Hur, D. (2007). Bipolarity in reactions to operational 'constraints': OM bugs under an OB lens. *Journal of Operations Management*, *25*(1), 1–13.

Bendoly, E., Croson, R., Goncalves, P., & Schultz, K. (2010a). Bodies of knowledge for research in behavioral operations. *Production and Operations Management*, *19*(4), 434–452.

Bendoly, E., Thomas, D., & Capra, M. (2010b). Multilevel social dynamics considerations for project management decision makers: Antecedents and implications of group member tie development. *Decision Sciences*, *41*(3), 459–490.

Boyer, K. K., & Verma, R. (2000). Multiple raters in survey-based operations management research: A review and tutorial. *Production and Operations Management*, *9*(2), 128–140.

Boyett, J. H., & Conn, H. P. (1992). *Workplace 2000: The revolution reshaping American business*. New York: Plume Books.

Brockman, B. K., Rawlston, M. E., Jones, M. A., & Halstead, D. (2010). An exploratory model of interpersonal cohesiveness in new product development teams. *Journal of Product Innovation Management*, *27*(2), 201–219.

Calder, B. J. (1977). An attribution theory of leadership. In B. M. Staw & G. R. Salancik (Eds.), *New directions in organizational behavior*. Chicago: St. Clair, 179–204.

Carson, J. B., Tesluk, P. E., & Marrone, J. A. (2007). Shared leadership in teams: An investigation of antecedent conditions and performance. *Academy of Management Journal*, *50*(5), 1217–1234.

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, *83*(2), 234–246.

Chen, C. K., Tsai, C. H., & Shu, K. C. (2009). An exploratory study for groupthink research to enhance group decision quality. *Journal of Quality*, *16*(2), 137–152.

Cohen, A., Doveh, E., & Eick, U. (2001). Statistical properties of the $r_{WG(J)}$ index of agreement. *Psychological Methods*, *6*(3), 297–310.

Cohen, S. G., & Bailey, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management*, *23*(3), 239–290.

Croson, R., Schultz, K., Siemsen, E., & Yeo, M. L. (2013). Behavioral operations: The state of the field. *Journal of Operations Management*, *31*(1), 1–5.

De Dreu, C. K. (2006). When too little or too much hurts: Evidence for a curvilinear relationship between task conflict and innovation in teams. *Journal of management*, *32*(1), 83–107.

De Dreu, C. K., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. *Journal of Applied Psychology*, *88*(4), 741.

DeNisi, A. S., Hitt, M. A., & Jackson, S. E. (2003). The knowledge-based approach to sustainable competitive advantage. In S. E. Jackson, M. A. Hitt, & A. S. DeNisi (Eds.), *Managing knowledge for sustained competitive advantage*. San Francisco: Jossey-Bass, pp. 3–33.

Dennis, A. R., Rennecker, J. A., & Hansen, S. (2010). Invisible whispering: Restructuring collaborative decision making with instant messaging. *Decision Sciences*, *41*(4), 845–886.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology*, *51*(3), 629–636.

Dong, Y., Zhang, H., & Herrera-Viedma, E. (2016). Integrating experts′ weights generated dynamically into the consensus reaching process and its applications in managing non-cooperative behaviors. *Decision Support Systems*, *84*, 1–15.

Druskat, V. U., & Wheeler, J. V. (2003). Managing from the boundary: The effective leadership of self-managing work teams. *Academy of Management Journal*, *46*(4), 435–457.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*(2), 350–383.

Edmondson, A. C., Bohmer, R. M., & Pisano, G. P. (2001). Disrupted routines: Team learning and new technology implementation in hospitals. *Administrative Science Quarterly*, *46*(4), 685–716.

Ensley, M. D., Hmieleski, K. M., & Pearce, C. L. (2006). The importance of vertical and shared leadership within new venture top management teams: Implications for the performance of startups. *The Leadership Quarterly*, *17*(3), 217–231.

Esser, J. K. (1995). *Groupthink from Pearl Harbor to the Challenger: Failure of decision-making groups*. Beaumont, TX: Lamar University.

Esser, J. K. (1998). Alive and well after 25 years: A review of groupthink research. *Organizational Behavior and Human Decision Processes*, *73*(2), 116–141.

Freedman, D. A. (1991). Statistical models and shoe leather. *Sociological Methodology*, *21*(2), 291–313.

Gibb, C. A. (1954). *Leadership*. In G. Lindzey (Ed.), *Handbook of social psychology*. Reading, UK: Addison-Wesley, 877–917.

Hackman, J. R., Walton, R. E., & Goodman, P. S. (1986). Leading groups in organizations. In *Designing effective work groups*. San Francisco: Jossey-Bass, 72–199.

Hamilton, B. H., & Nickerson, J. A. (2003). Correcting for endogeneity in strategic management research. *Strategic Organization*, *1*(1), 51–78.

Hammer, M., & Champy, J. (1993). *Reengineering the corporation: A manifesto for business revolution*. New York: Harper Business.

Harvey, J. B. (1974). The abilene paradox: The management of agreement. *Organizational Dynamics*, *17*(1), 63–80.

Harvey, M., Novicevic, M. M., Buckley, M. R., & Halbesleben, J. R. (2004). The abilene paradox after thirty years: A global perspective. *Organizational Dynamics*, *33*(2), 215–226.

Herek, G. M., Janis, I. L., & Huth, P. (1987). Decision making during international crises: Is quality of process related to outcome? *Journal of Conflict Resolution*, *31*(2), 203–226.

Hoegl, M., Parboteeah, K. P., & Munson, C. L. (2003). Team-level antecedents of individuals' knowledge networks. *Decision Sciences*, *34*(4), 741–770.

Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. *Journal of management*, *23*(6), 723–744.

Huber, G. P., & Power, D. J. (1985). Retrospective reports of strategic-level managers: Guidelines for increasing their accuracy. *Strategic Management Journal*, *6*(2), 171–180.

Huckman, R. S., Staats, B. R., & Upton, D. M. (2009). Team familiarity, role experience, and performance: Evidence from Indian software services. *Management Science*, *55*(1), 85–100.

Janis, I. L. (1982). *Groupthink*. Boston: Houghton-Mifflin.

Janis, I. L., & Mann, L. (1977). *Decision making: A psychological analysis of conflict, choice, and commitment*. New York: The Free Press.

Jehn, K. A., & Mannix, E. A. (2001). The dynamic nature of conflict: A longitudinal study of intragroup conflict and group performance. *Academy of Management Journal*, *44*(2), 238–251.

Jong, A. D., & De Ruyter, K. (2004). Adaptive versus proactive behavior in service recovery: The role of self-managing teams. *Decision Sciences*, *35*(3), 457–491.

Katz, D., & Kahn, R. L. (1978). *The social psychology of organizations*(2nd ed). New York: Wiley.

Ketokivi, M. A., & Schroeder, R. G. (2004). Perceptual measures of performance: Fact or fiction?. *Journal of Operations Management*, *22*(3), 247–264.

Kim, Y. (2001). A comparative study of the" Abilene Paradox" and "Groupthink." *Public Administration Quarterly*, *25*(2), 168–189.

Klein, K. J., & Kozlowski, S. W. J. (2000). *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. San Francisco: Jossey-Bass.

Knight, D., Pearce, C. L., Smith, K. G., Olian, J. D., Sims, H. P., Smith, K. A., & Flood, P. (1999). Top management team diversity, group process, and strategic consensus. *Strategic Management Journal*, 445–465.

Kozlowski, S. W., & Bell, B. S. (2003). Work groups and teams in organizations. Handbook of psychology. In W. C. Borman & D. R. Ilgen (Eds.), *Comprehensive handbook of psychology: Industrial and organizational psychology*. New York: Wiley, 333–375.

Kozlowski, W. J., & Hattrup, K. (1992). A disagreement about within-group agreement: Disentangling issues of consistency versus consensus. *Journal of Applied Psychology*, *77*(2), 161–167.

Kreft, I. G. G., & de Leeuw, J. (1998). *Introducing Multilevel Modeling*. Newbury: Park Sage.

Lance, C. (1988). Residual centering, exploratory and confirmatory moderator analysis, and decomposition of effects in path models containing interactions. *Applied Psychological Measurement*, *12*(2), 163–175.

Lapré, M. A., & Tsikriktsis, N. (2006). Organizational learning curves for customer dissatisfaction: Heterogeneity across airlines. *Management Science*, *52*(3), 352–366.

Leana, C. R. (1985). A partial test of Janis's groupthink model: Effects of group cohesiveness and leader behavior on defective decision making. *Journal of Management*, *11*(1), 5–18.

Lord, R. G., & Maher, K. J. (2002). *Leadership and information processing: Linking perceptions and performance*, London: Routledge.

Lowry, P. B., Roberts, T., Douglas, L. D., & Marakas, G. M. (2007). Toward building self-sustaining groups in PCR-based tasks through implicit coordination:

The case of heuristic evaluation. *Journal of the Association for Information Systems*, *10*(3), 170–195.

Manz, C. C., & Sims, H. P. (1982). The potential for “groupthink” in autonomous work groups. *Human Relations*, *35*(9), 773–784.

Manz, C. C., & Sims, H. P., Jr (1987). Leading workers to lead themselves: The external leadership of self-managing work teams. *Administrative Science Quarterly*, *32*(1), 106–129.

Mao, K. Z. (2004). Orthogonal forward selection and backward elimination algorithms for feature subset selection. *IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics*, *34*(1), 629–634.

Maznevsky, M. (1994). Understanding our differences: performance in decision-making groups with diverse members. *Human Relations*, *47*(5), 531–552.

McAvoy, J., & Butler, T. (2009). The role of project management in ineffective decision making within Agile software development projects. *European Journal of Information Systems*, *18*(4), 372–383.

McFadden, K. L., Henagan, S. C., & Gowen, C. R. (2009). The patient safety chain: Transformational leadership’s effect on patient safety culture, initiatives, and outcomes. *Journal of Operations Management*, *27*(5), 390–404.

Meindl, J. (1993). Reinventing leadership: A radical, social psychological approach. In J. K. Murnighan (Ed.), *Social psychology in organizations*. Englewood Cliffs: Prentice Hall.

Mihalache, O. R., Jansen, J. J., Van Den Bosch, F. A., & Volberda, H. W. (2012). Offshoring and firm innovation: The moderating role of top management team attributes. *Strategic Management Journal*, *33*(13), 1480–1498.

Mishra, A., Chandrasekaran, A., & MacCormack, A. (2015). Collaboration in multi-partner R&D projects: The impact of partnering scale and scope. *Journal of Operations Management*, *33*, 1–14.

Mohammed, S., & Ringseis, E. (2001). Cognitive diversity and consensus in group decision making: The role of inputs, processes, and outcomes. *Organizational Behavior and Human Decision Processes*, *85*(2), 310–335.

Montoya-Weiss, M. M., Massey, A. P., & Song, M. (2001). Getting it together: Temporal coordination and conflict management in global virtual teams. *Academy of Management Journal*, *44*(6), 1251–1262.

Moorhead, G., & Montanari, J. R. (1986). An empirical investigation of the groupthink phenomenon. *Human Relations*, *39*(5), 399–410.

Moreland, R. L., Levine, J. M., & Wingert, M. L. (2013). Creating the ideal group: Composition effects at work. In E. Witte & J. H. Davis (Eds.), *Understanding group behavior* (Vol. 2). New York: Psychology Press, Taylor & Francis Group, 11–35.

Morrison, E. W. (2011). Employee voice behavior: Integration and directions for future research. *Academy of Management Annals*, *5*(1), 373–412.

Neck, C. P., & Moorhead, G. (1995). Groupthink remodeled: The importance of leadership, time pressure, and methodical decision-making procedures. *Human Relations*, *48*(5), 537–557.

Neubert, M. J. (1999). Too much of a good thing or the more the merrier? Exploring the dispersion and gender composition of informal leadership in manufacturing teams. *Small Group Research*, *30*(5), 635–646.

O'Neill, T. A., & McLarnon, M. M. (2017). Optimizing team conflict dynamics for high performance teamwork. *Human Resource Management Review*, https://doi.org/10.1016/j.hrmr.2017.06.002

Park, W. W. (2000). A comprehensive empirical investigation of the relationships among variables of the groupthink model. *Journal of Organizational Behavior*, *21*(8), 873–887.

Pearce, C. L., & Sims, H. P. (2002). The relative influence of vertical vs. shared leadership on the longitudinal effectiveness of change management teams. *Group Dynamics: Theory, Research, and Practice*, *6*(2), 172–197.

Perry-Smith, J. E., & Shalley, C. E. (2003). The social side of creativity: A static and dynamic social network perspective. *Academy of Management Review*, *28*(1), 89–106.

Pfeffer, J. (1977). The ambiguity of leadership. *Academy of Management Review*, *2*(1), 104–114.

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2004). *HLM 6 hierarchical and nonlinear modeling*. Chicago: Scientific Software International.

Ren, Y., Kiesler, S., & Fussell, S. R. (2008). Multiple group coordination in complex and dynamic task environments: Interruptions, coping mechanisms, and technology recommendations. *Journal of Management Information Systems*, *25*(1), 105–130.

Riccobono, F., Bruccoleri, M., & Größler, A. (2016). Groupthink and project performance: The influence of personal traits and interpersonal ties. *Production and Operations Management*, *25*(4), 609–629.

Roberts, M. R., & Whited, T. M. (2013). Endogeneity in empirical corporate finance. In In G. M. Constantinides, M. Harris, & R. M. Stulz (Eds.), *Handbook of the Economics of Finance* (Vol. *2A*). Amsterdam: Elsevier, 493–572.

Roch, S. G. (2007). Why convene rater teams: An investigation of the benefits of anticipated discussion, consensus, and rater motivation. *Organizational Behavior and Human Decision Processes*, *104*(1), 14–29.

Rotter, J. B. (1982). *The development and applications of social learning theory*. New York: Praeger.

Schmidt, J. B., Montoya-Weiss, M. M., & Massey, A. P. (2001). New product development decision-making effectiveness: Comparing individuals, face-to-face teams, and virtual teams. *Decision Sciences*, *32*(4), 575–600.

Schweiger, D. M., Sandberg, W. R., & Ragan, J. W. (1986). Group approaches for improving strategic decision making: A comparative analysis of dialectical inquiry, devil's advocacy, and consensus. *Academy of Management Journal*, *29*(1), 51–71.

Shaw, M. E. (1964). Communication networks. In L. Nerkowitz (Ed.), *Advances in experimental social psychology* (Vol. *1*). New York: Academic, 111–147.

Sillito Walker, S. D., & Bonner, B. L. (2018). The effects of differing knowledge transfer strategies on group decision making and performance. *Journal of Behavioral Decision Making*, *31*, 115–126.

Simons, T. L., & Peterson, R. S. (2000). Task Conflict and Relationship Conflict in Top Management Teams: The Pivotal Role of Intragroup Trust. Retrieved [July 2017], from Cornell University, School of Hospitality Administration site: Retrieved from https://scholarship.sha.cornell.edu/articles/719

Sims, R. R. (1994). *Ethics and organizational decision making: A call for renewal*. Greenwood: Publishing Group.

Sinclair, A. (1992). The tyranny of a team ideology. *Organization studies*, *13*(4), 611–626.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, *23*(4), 323–355.

Sivasubramaniam, N., Murry, W. D., Avolio, B. J., & Jung, D. I. (2002). A longitudinal model of the effects of team leadership and group potency on group performance. *Group & Organization Management*, *27*(1), 66–96.

Sparrowe, R. T., Liden, R. C., Wayne, S. J., & Kraimer, M. L. (2001). Social networks and the performance of individuals and groups. *Academy of management journal*, *44*(2), 316–325.

Stewart, G. L., & Manz, C. C. (1995). Leadership for self-managing work teams: A typology and integrative model. *Human relations*, *48*(7), 747–770.

Stogdill, R. M. (1948). Personal factors associated with leadership: A survey of the literature. *Journal of Psychology*, *25*(1), 35–71.

Swaab, R. I., Phillips, K. W., & Schaerer, M. (2016). Secret conversation opportunities facilitate minority influence in virtual groups: The influence on majority power, information processing, and decision quality. *Organizational Behavior and Human Decision Processes*, *133*, 17–32.

Tsai, M. H., & Bendersky, C. (2015). The pursuit of information sharing: Expressing task conflicts as debates vs. disagreements increases perceived receptivity to dissenting opinions in groups. *Organization Science*, *27*(1), 141–156.

Wellman, B., & Berkowitz, S. D. (1988). *Social structures: A network approach*. New York: Cambridge University Press.

Williams, K. Y., & O'Reilly, C. A. (1998). Demography and diversity in organizations: A review of 40 years of research. *Research in Organizational Behavior*, *20*(20), 77–140.

Zaccaro, S. J., Rittman, A. L., & Marks, M. A. (2002). Team leadership. *The Leadership Quarterly*, *12*(4), 451–483.

Zhang, D., Lowry, P. B., Zhou, L., & Fu, X. (2007). The impact of individualism-collectivism, social presence, and group diversity on group decision making under majority influence. *Journal of Management Information Systems*, *23*(4), 53–80.

Zhou, K. Z., & Li, C. B. (2012). How knowledge affects radical innovation: Knowledge base, market knowledge acquisition, and internal knowledge sharing. *Strategic Management Journal*, *33*(9), 1090–1102.

Dr. Manfredi Bruccoleri is associate professor of operations management and business process management and teaches for bachelor and master students of the University of Palermo. He holds a doctoral degree in Production Engineering and in 2001 he was visiting scholar at the School of Engineering of the University of Michigan (Ann Arbor). He collaborates with companies from different manufacturing and service industries for both teaching and research projects and coordinated several business process reengineering projects in different Italian companies and Public Administration. His research interests focus on operations management, behavioral operations, supply chain management, reverse logistics, business process management, social responsibility. He is member of the Editorial Review Board of International Journal of Operations and Production Management.

Dr. Francesca Riccobono is a post-doc research scientist at the University of Palermo. From the same University she received her bachelor and master degrees in Management Engineering and a PhD in Production Engineering in 2012. She was visiting master student at the Enterprise Research Centre of the University of Limerick (Ireland) in 2008 and visiting PhD student at Columbia Business school (New York, NY) in 2011. Her research interest mainly concerns networking strategy, innovation and R&D management and behavioral operations. She is member of the European Operations Management Association (EUROMA) and the Italian Association of Engineering Management (AiIG).

Dr. Andreas Größler is full professor of operations management at University of Stuttgart (Germany) where he currently is also managing director of the Institute of Business Administration. He teaches operations strategy and behavioral operations management courses at the undergraduate, master, and doctoral level. His research interests lie in the fields of sustainable operations management, in applying system dynamics modelling and simulation in business and not-for-profit organizations, and in investigating individual and organizational dynamic decision-making. Dr. Größler was a managing editor of System Dynamics Review and is an associate editor of International Journal of Operations and Production Management and Operations Management Research.