Linking Team Composition to Team Performance: An Application to Postdisaster Debris Removal Operations

David Mendonça, *Senior Member, IEEE*, James D. Brooks, *Member, IEEE*, and Martha Grabowski

***Abstract*—This paper considers how changes in team composition (such as the number and rate of turnover of team members) are linked to team performance, as assessed in terms of efficiency, effectiveness, and equality (i.e., distribution of effort). Study data are taken from a large-scale, postdisaster debris removal operation in the USA, collected through existing transaction-level data logging systems. The data enable a detailed (and objective) examination of team performance, thus overcoming many shortcomings of retrospective methods such as questionnaires. The results show that the increased turnover diminishes performance along all dimensions, while an increased team size contributes to effectiveness but reduces equality. Implications of this study for theory and future empirical work are both discussed.**

***Index Terms*—Human factors, man–machine systems, system performance, teamwork.**

I. Introduction

In disaster response operations, teams may be dispatched to the field to enable rapid and effective mitigation of immediate postdisaster effects. These teams—whether in search and rescue, firefighting, police work, or utility restoration—may range in size and be comprised of individuals with varying degrees of experience in working together. However, this variation in team composition has been theorized as both as a threat and a potential enhancement to team performance [1], [30]. For example, as discussed by Levine *et al*. [30], changes in team size and working history may have positive or negative effects on performance, depending on whether the costs of these changes (e.g., in terms of coordination) outweigh the benefits of working together [1].

Team composition and performance are investigated here through an analysis of field data collected in the domain of postdisaster debris removal operations: a highly distributed, increasingly expensive activity in which potentially hundreds of teams are employed, with work stretching over many months. The organization that directs these operations is best characterized as *team-centric*: employing hundreds of teams with thousands of team members. Changes to team composition occur often, and are reflected in fluctuations in the number and working history of team members over time. Because the cost of debris removal is largely reimbursed through public funds, detailed documentation is kept for cost accounting purposes (e.g., in determining where, when, and by whom debris was transported from one location to another)—thus creating a machine-readable dataset covering various aspects of teamwork.

This paper takes a three-pronged approach to investigate the differential effects of team composition on team performance. First, it follows recent research in articulating how process-level measures may contribute to research on teamwork in organizations. Second, drawing upon related research in decision making studies e.g., [7], [14], it argues for unpacking the concept of performance into three constituent dimensions—effectiveness, efficiency, and equality (i.e., equal distribution of effort)—to attempt to reconcile prior conflicting results. Third, it investigates questions of team composition and performance with data collected via largely automated methods in field settings, thereby reducing some threats associated with nonresponse and response bias.

The paper is organized as follows. Theoretical background is provided first to motivate the main research questions in Section II, followed by the methodology in Section III, results and discussion in Section IV, and conclusions in Section V.

II. Theoretical Background

Teams are now recognized as essential to the work of many organizations. As noted elsewhere, definitions of teams are varied in scope and specificity. In line with much recent and prior work e.g., [54], a team is here understood as "two or more persons with a common goal that requires interdependence and adaptive functioning" [43]. For comparison, "work teams"—a type of team closely akin to that considered here—are described by Kozlowski and Bell ([28], as cited in [54]):

> "collectives who exist to perform organizationally relevant tasks, share one or more common goals, interact socially, exhibit task interdependences, maintain and manage boundaries, and are embedded in an organizational context that sets boundaries, constrains the team, and influences exchanges with other units in the broader entity."

Manuscript received April 23, 2013; revised October 28, 2013 and January 16, 2014; accepted January 17, 2014. Date of publication March 3, 2014; date of current version May 13, 2014. This work is supported in part by the National Science Foundation under Grant 1313589. This paper was recommended by Associate Editor C. Burns.

D. Mendonça and J. D. Brooks are with the Department of Industrial and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: mendod@rpi.edu; brookj7@rpi.edu).

M. Grabowski is with the Department of Industrial and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA and with the Department of Business Administration and the Information Systems Program, LeMoyne College, Syracuse, NY 13212 USA (e-mail: grabowsk@lemoyne.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/THMS.2014.2306198

As suggested by this definition, teams within organizations may be deployed to enable greater responsiveness to changing conditions in the field, where they may operate in "unstable environments" in which they must "adapt and respond quickly to changing conditions" [21]. Indeed, definitions of teams (or groups) at times encompass characteristics both of team members and of the task(s) that the team performs. In Goodman *et al.*'s work on mining crews (as well as in related applications such as logging), the value of a product is created at one level of the organization, with other levels providing coordination and management. In contrast to teams performing such a "production" work are those performing "knowledge work," where the value of a product is created as it moves through various levels of the organization, as in product design and software engineering, e.g., [8].

Issues of interdependence, goal-setting, and adaptation may be seen in production and knowledge settings. For example, extraction of ore by mining crews requires a series of coordinated physical steps which cannot be accomplished by a single individual; similarly, computer programmers performing the "virtual" work of programming may subdivide and sequence their work [37].

More recent work has separated team from task-level elements in defining teams [34], [54]. Task interdependence—the extent to which the outcomes of team members are influenced by the actions of others [54]—has been identified as a key factor for discriminating amongst types of teams [16], [28]. Indeed, recent work has emphasized the salience of task interdependence e.g., [11] and [34], and of the broader need for team categorization systems based on conceptually distinct and theoretical dimensions (see [54] for an example). It is expected that the further development and validation of such classification systems will begin to produce a more consistent set of methods for classifying teams [49]. The term *team* is used in the remainder of this theoretical background, and the actual teams under study discussed in Section III.

Team adaptation and response may occur through a variety of mechanisms [18], including changes in the number and working history of team members (here denoted team size and team fluidity, respectively). Much prior work on team composition has focused on developing measures that reflect the characteristics of individual members of the team, such as the gender and cognitive ability of members, as well as their tenure within the organization. More recent work has expanded the team composition construct to include team-level measures, such as team fluidity and team size. Team fluidity, commonly understood as change in membership [46], is a multifaceted term, encompassing structural and procedural changes as well as changes to personnel [46]. Indeed, if the extent of fluidity with respect to personnel changes is extremely high, all that may remain of a team after some time is its name. Similarly, teams may be formed, dissolved, and reformed over such small time horizons that it is more appropriate to discuss fluidity in terms of average level of familiarity between pairs of members [13]. Team size typically refers to the number of individuals that comprise the team, although frequently—and particularly in the presence of team fluidity—this may be a difficult measure to estimate. For example, team members may participate only briefly in team work, and thus feel themselves to be only weakly associated with other team members. Because fluidity is an inherently dynamic concept, various researchers have argued for collecting data that attempt to capture "moment to moment interteam and intrateam interaction" [44], including how these interactions unfold over longer time scales [12].

A. Team Fluidity and Team Performance

Changes in team fluidity—whether toward greater rigidity or greater flexibility [46]—are expected to produce significant effects on team performance. However, there has been a considerable debate in the research literature over the direction of these effects and the conditions under which they occur. As discussed by Bell and Kozlowski [4], while fluidity is expected to be substantially greater in teams within virtual organizations [50]—thus permitting a more flexible organizational response—high fluidity can also create "conflicts and ambiguity" [40]. According to Levine *et al.* [30], membership change may have positive or negative effects on performance, depending on whether this change outweighs the benefits of working together [1]. Consequently, changes in team composition have been theorized both as a potential threat and potential enhancement to team performance [1], [30] and, by extension, to the organization as a whole. Furthermore, Saunders and Ahuja [45] found that satisfaction is less important to the members of temporary teams. Cohesion of a team grows over time [24], and thus is adversely affected by team fluidity [3].

Team performance is here viewed as a construct with multiple dimensions—such as efficiency, effectiveness, and equality—associated with the outcomes of team decision making (sometimes denoted "outcome performance" [5]). A fundamental result of this study is the presence of tradeoffs across different dimensions of performance. For example, past related work has considered effectiveness (i.e., the overall quality or amount of work accomplished by the team) versus efficiency (i.e., output per unit of effort, as in [3]), efficiency versus equality (i.e., fairness of distribution of effort) [6], or all three dimensions [15]. Alternative conceptualizations of team performance of course exist: for example, effectiveness itself has been cast as an encompassing term, rather than as a dimension of performance, which reflects team performance (here, effectiveness), team process, and individual member satisfaction [23], [52].

Team performance may depend more on individual expertise for the performance of knowledge work (where routines are evolving and new ideas are essential) than for the performance of production work (where the focus is on reliable performance of routine and structured tasks). Because lateral or creative thinking is required in knowledge work, new team members may bring new ideas, thus increasing team capabilities [30]. In the production work, on the other hand, fluidity may decrease productivity (i.e., effectiveness), as new members strive to learn already established (and non-mutable) patterns of work. These observations lead to the following related research propositions.

Proposition 1.1: Increased fluidity is expected to decrease performance.

Proposition 1.2: However, the impact of increased fluidity is expected to diminish over time, as roles stabilize and fluidity becomes institutionalized.

As teams increase in size, their production capacity—and thus their effectiveness—is expected to increase: indeed, the decision to add members to a team is likely to be a response to increased workload demands. There is a substantial literature linking perceived workload and compensation equality to other outcomes such as absenteeism, satisfaction, and work quality [9], [26]. For example, teams performing a production task have been found to exhibit decreased differences in individual performance (i.e., greater equality) as team size increases [48]. Finally, increased team size is also expected to increase complexities associated with coordination of teamwork [47]. Further, work on social loafing has indicated that individuals decrease their effort as the number of coworkers is increased [29], thus reducing team efficiency. These observations on the effect of team size on performance are captured in the following research proposition.

Proposition 2: Increased team size will be associated with increased effectiveness and equality, but lower efficiency.

As has been noted elsewhere [14], performance tradeoffs are to be expected between the dimensions of effectiveness, efficiency, and equality. For example, decisions about school staffing based purely on an effectiveness criterion may result in hiring more teachers, those based on efficiency may result in hiring of fewer teachers, and those based on equality would distribute teachers evenly among schools. Results on the relationship between team fluidity and team performance have been inconsistent [20], perhaps due to the rarity with which these tradeoffs are explicitly modeled. In one notable exception, teams which know they are temporary tend to weight effectiveness more heavily than efficiency [45], leading to the final research proposition.

Proposition 3: Increased fluidity will be associated with greater effectiveness relative to efficiency.

Fig. 1. Activities in the Debris Removal Mission.

III. Methodology

This section provides a brief description of the study domain, and then explains the methods and data used to address the previous propositions.

A. Domain Background

Debris removal is a critical but understudied phase in disaster response and recovery [31], [39], [51]. As has been made clear from many recent and historical disasters, debris must be removed from rights-of-way and, occasionally, from private property in order for rebuilding to begin. In the United States, large-scale debris removal operations that cover significant portions of one or many states are coordinated by the US Army Corps of Engineers (USACE)—under contract from the Department of Homeland Security—which facilitates management and monitoring of the contractors and subcontractors who remove debris and deliver it to temporary sites for disposal.

The debris removal mission is characterized by team-centered work, as shown in Fig. 1. The debris field (panel 1) consists of a potentially wide variety of material distributed over the affected area. Debris on private property is typically pushed to the curbside by property owners, where it is segregated and loaded onto trucks and trailers at designated pickup points by debris removal teams (DRTs – panel 2). On a given day, a DRT typically performs pickup activities over a demarcated area (e.g., a small number of city blocks) to which it has been assigned.

Teamwork is integral to the job functions of DRTs. The work of teams is directed toward the common goal [28] of clearance of debris. At the pickup site, DRT personnel tasked with loading must work together to reduce the debris pile at curbside (see Fig. 1, panel 2). At a most basic level, *task interdependence* [34] is determined by the physical limitations of personnel and equipment, as well as spatial limitations of the pickup site. At curbside, crew members must sometimes work collectively to move heavy debris by hand to where it can be grasped by a mechanical loader and placed on a waiting truck. Loading continues until loaders determine when the trucks will leave the pickup site (i.e., not all trucks leave 100% full). Additionally, when trucks exit or enter the pickup site, DRT personnel must adjust work processes to accommodate them.

A second aspect of teamwork, *adaptation* either by individual members or the team itself, is reflected in loading decisions by team members about 1) which trucks should receive which debris, 2) which debris to load at which time (some debris is easier and/or faster to load than other debris, and multiple passes may necessary in order to clear it), and 3) when a load is considered complete. Routines are sufficiently interlocking that supervisors are assigned to monitor and manage teams and their workflow, as well as to maintain safety. Thus, while drivers' (and other team members') decisions may engender minor variation in loading and unloading performance, these minor variations must be scaled to the sometimes considerable number of loads hauled.

Upon leaving the loading site, trucks travel the road network to deliver loads to a temporary debris storage and reduction site (TDSR, panel 3), where the size of the load is estimated visually as a percentage of the truck's total capacity by a human operator. A guarantee of payment (called a load ticket) is delivered to the truck driver, who then travels the road network to rejoin the DRT at the pickup point.

The trucks used to load and haul debris are at the core of each DRT. Truck size constrains the volume of debris that can be hauled from the pickup site, and the choices that their drivers make regarding routing to and from TDSRs help determine the number of loads that can be hauled in a given day. Discussions with field personnel also suggest that staffing and other equipment decisions are driven by the number and mix of trucks available for assignment to a pickup site. Moreover, any trucks assigned to the mission must be inspected and certified to work—a process that is completed for each mission. Trucks are typically driven by their owner for the duration of the mission. The number and working history of trucks in a given DRT can vary according to the demands of the mission.

As suggested by the foregoing discussion, teams are embedded within the organization [28] and central to its mission. At the organizational level, dispatchers assign personnel and equipment to teams (and also specify pick up and drop off points). Operational details of loading and unloading activities are left to the teams themselves, subject to safety and environmental considerations (e.g., adverse weather, disabled equipment) and any additional organizational-level constraints (e.g., when to start and stop work).

However, there is no designed interdependence—and only minimal interaction—*among* teams: a DRT works at curbside without the collaboration or interference of other DRTs; travel to the TDSR may be affected by roadway congestion (in part a consequence of the activities of other DRTs); and drop-off activities may be delayed by the presence of other trucks queued for service at the TDSR. A weak form of between-team interaction (aside from queuing) is the reassignment of trucks from one team to another. Consequently, while the organization is team-centric, teams operate nearly independently, in contrast with teams in multiteam systems, where they "interface directly and interdependently" [36]. Accordingly, a team-level (as opposed to organizational-level) analysis is undertaken here.

Recent advances in information technology for tracking debris removal and disposal, support a more detailed and precise examination of DRT performance. The load ticket functions as an accounting record of truck—and therefore DRT—activity, providing proof that a given load has been hauled from a pickup site and delivered to a TDSR. In recent years, the load ticket has become electronic: automated debris management systems provide detailed data on a load's history, including time and location of pick-up and drop-off activities, load size, and various other attributes. This work employs these data to develop measures of key study constructs.

B. Characteristics of Teams: Study Measures

This section discusses and summarizes key study variables associated with team composition and team performance. Process-level measures of teamwork, particularly objective ones, can provide the level of detail necessary to investigate the dynamics of team composition at an appropriate scale. Indeed, recent work has called for [42], and in some cases illustrated e.g., [13], more extensive use of study measures of team- and organization-level processes, particularly in field settings [42]. However, field work on teams within organizations has been hampered by challenges associated with the sufficiency of researchers' resources and the complexity of field settings [22], [41]. Data must be available (e.g., not all organizations track team composition and performance in ways that are theoretically relevant), and must be framed and interpreted in relation to the organizational processes from which they have been extracted. As has been stated elsewhere [13], research along these lines may serve as a first step toward developing models that are built from fairly low-level data that explicitly engage team- and organizational-level dynamics [25].

1) Team Composition: As discussed previously, the work of DRTs is constrained by the number and hauling capacity of trucks operating within it. Because each subcontractor is reimbursed for their work based on the total cubic yards hauled per day, their earnings are a function of the performance of trucks operating within their teams. Although there are other elements of the team, such as loading (i.e., nontruck driving) personnel and other nontruck equipment, that contribute to the work, debris cannot be loaded and delivered without trucks. Consequently, the composition of DRTs is here expressed in terms of the composition of trucks that are integrated within it.

Team composition is reconstructed from the transactional (i.e., load ticket) data by taking advantage of data reporting protocols in the field. Each DRT has one and only one Quality Control (QC) officer assigned to it, stationed at the pickup point. Responsibilities of the QC include issuing load tickets, monitoring safety, verifying eligibility of debris, ensuring compliance of team activities, and communicating with the DRT's on-site supervisor. In meeting these responsibilities, QCs are required to maintain line-of-sight observation of the DRT. Due to this task structure, team membership may be inferred from the load ticket data by grouping tickets which share a common QC attribute on a given day. The *size* of a DRT is therefore expressed in terms of the number of trucks operating within it, while the *fluidity* of a DRT is expressed in terms of the working history between trucks (recall that truck owners operate their trucks for the duration of their involvement in the mission). Both of these team characteristics are now discussed in detail.

Team Fluidity: The approach to addressing the issue of fluidity at the team level is motivated by the recommendation that more "process-oriented" measures be used to identify roles within teams. A process-oriented approach may include, for example, examining "distinguishable roles" and "patterns of expected behavior and norms" [35]. Because team fluidity is by definition a dynamic phenomenon, its study therefore calls for "unobtrusive and real-time measures of team performance that can be practically implemented, especially in the field" [42], [53] in order to capture "moment to moment interteam and intrateam interaction" [44]. Consistent with prior work [13], team fluidity is here defined as a function of the working history among all dyads in the group. The measure is composed by drawing on the definition of *familiarity* [13] shown in (1). The familiarity measure may be understood as expressing average shared working experience among all pairs of team members. For a team with n members, h_{ij} captures number of days that each possible pair of trucks (i, j) in a team has worked together. The sum is then normalized by the number of possible pairs in the team [the leading coefficient in (1)]. The familiarity of team k, denoted by f_k, is then given by

$$f_k = \frac{2}{n\,(n-1)} \left(\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} h_{ij} \right). \tag{1}$$

The following example illustrates the measure. Consider five individual trucks (for simplicity of illustration, here denoted as A–E) that enter the system, none having worked together previously. For the first two days, A–C work together as one team (team 1 = $\{$A–C$\}$), and D and E as another team (team 2 = $\{$D, E$\}$). The history of each possible dyad in the team is calculated at the start of each day. At the start of day two, all dyads in each team have a familiarity of unity from the prior day's experience. For example, for team 1, the sum of the working histories is 3, and the value of the leading coefficient is 2/(3 × 2) = 1/3, yielding a familiarity of unity. Similarly, the familiarity for team 2 is also unity. Now suppose that on day 3 the teams are changed to be team 1 = $\{$A, C, E$\}$ and team 2 = $\{$B, D$\}$. The value of h_{AC}, the cumulative working history for dyad (A, C), is now 2, while $h_{AE} = h_{CE} = 0$ since neither of these pairs has worked together previously. As a result, the average pair-wise familiarity for team 1 is $f_1 = (2 + 0 + 0)/3 =$ 2/3. Similarly, for team 2, B and D have not yet worked together, so $h_{BD} = 0$, and thus $f_2 = 0$.

With this formulation, the maximum familiarity possible grows with time (e.g., at 10 days into the mission, it is 10; while at 20, it is 20). Familiarity is therefore further scaled by the total number of prior days d since the onset of the mission (i.e., two in the example) to yield a measure denoted *relative familiarity*. In this example, the value of relative familiarity for team 1 on the third day is $f_1/2 = 0.33$, and for team 2 it is 0. This measure is thus bounded between 0 and 1, inclusive.

Fluidity of team k, denoted F_k, may then be viewed as the inverse of relative familiarity, and is calculated as $F_k = 1 - f_k/d$. In the example, Team 1 then has a fluidity of $F_1 = 1 - 0.33 = 0.67$, while Team 2 has a fluidity of $F_2 = 1 - 0 = 1$. In this way, the teams whose members have no prior experience are given the maximum fluidity value while those having members with some prior history are given a lower fluidity value. A team of any size whose members have always worked together since the onset of the debris mission would receive a fluidity value of zero.

Team Size: As discussed previously, debris removal work depends critically on the hauling trucks within DRTs. Indeed, as confirmed in discussions with subject matter experts, the number of personnel working in a DRT is expected to correlate strongly with the number of trucks: an excess of trucks relative to personnel would lead to long wait times for trucks at loading zones, whereas a shortage of trucks would lead to lost worker hours as personnel perform under their capability. Team size is therefore expressed here as the number of trucks associated with a debris removal team.

2) Team Performance: Team performance is here expressed as a multidimensional construct [43], comprising efficiency, effectiveness, and equality. In line with Mathieu *et al.* [34], a blended or composite approach to the measurement of each dimension is used. Team-level measures are computed from measurements taken on individual team members [35], in light of actual performance incentives. Let the set of trucks in team k be denoted by $\mathcal{T}_k$ and let $\mathcal{H}_i$ be the set of loads hauled by truck i. Then, the total effective loads hauled by truck i for the day of interest l_i is given by

$$l_i = \sum_{m \in \mathcal{H}_i} p_m$$

where p_m is the percentage of the total truck volume for load m (i.e., $p_m \in [0, 1]$). This individual-level measure forms the basis for the team-level performance measures across the three dimensions of effectiveness, efficiency, and equality, as described next.

Effectiveness refers to a team's total output relative to a goal. In the case of a debris removal mission as a whole, effectiveness is assessed by tracking the total cubic yards delivered to TDSRs each day, with subcontractors compensated according to the volume of debris delivered by their teams. While there is no set goal for individual teams, the total debris delivered by the team may therefore be viewed as closely aligned to the mission level goal (particularly given the presence of an on-site supervisor at the pickup point). To account for differences in truck capacity, the effectiveness measure used is the number of effective loads hauled rather than cubic yards. This measure for team k is then given by

$$E_k = \sum_{i \in \mathcal{T}_k} l_i. \tag{2}$$

Efficiency refers to a team's production output relative to its capabilities. The capabilities of a DRT are fundamentally limited by the size of the team and the average haul distance to the TDSR. This dimension is therefore operationalized as the

TABLE I
NUMERICAL EXAMPLE: TRUCK DATA FOR A SINGLE TEAM

Truck (i)	*Effective Loads* (l_i)	*Avg. Distance* (d_i)
1	8.3	3.4
2	12.7	3.5
3	9.1	4.7

average number of effective load miles hauled per truck on a given day, as follows:

$$\eta_k = \sum_{i \in \mathcal{T}_k} \frac{l_i d_i}{n} \tag{3}$$

where d_i is the average haul distance for truck i and n is the team size (i.e., $n = |\mathcal{T}_k|$).

Equality refers to the similarity of individual team-level effort (i.e., equality of workload distribution across team members). For historical reasons, the inverse of this measure, denoted inequality, is often used. There are various inequality measures found in the literature [33]. Coulter's measure is used here [10], which is given generally by

$$I = \frac{1}{K}\sqrt{\sum_i \left[\frac{E_i}{\bar{E}} - \frac{A_i}{\bar{A}}\right]^2} \tag{4}$$

where E_i and A_i are the effect and relevant attribute for entity i, while $\bar{E}$ and $\bar{A}$ are the mean values across all entities. The scaling factor K can be used to provide a normalized measure (i.e., to ensure $I \in [0, 1]$). Equation (4) provides a measure of the match between the effect (e.g., work being done) and some group attribute (e.g., work to do), while penalizing higher levels of inequality, and providing a unit-less, normalized measure adhering to the accepted guidelines of scale invariance and the principle of transfers [32], [33]. Here, each i corresponds to a truck and the effect is the number of effective trips hauled. If uniform workload is desired (i.e., is perceived to be fair), then this measure for team k simply becomes

$$I_k = \frac{1}{\sqrt{n(n-1)}}\sqrt{\sum_{i \in \mathcal{T}_k}\left(\frac{l_i}{\bar{l}_k} - 1\right)^2}$$

where $\bar{l}_k$ is the number of effective loads hauled by the average truck in team k. The equality of a team is then the inverse of the inequality, given by

$$\mathcal{E}_k = 1 - I_k. \tag{5}$$

Numerical Example: Suppose trucks 1, 2, and 3 are working together on a particular day. As shown in Table I, truck 1 hauls 8.3 effective loads (suppose this is the result of 10 trips having the following load percentages, relative to the capacity of the truck: 0.8, 0.95, 0.9, 0.75 0.85, 0.7, 0.9, 0.75, 0.95, and 0.75) an average distance of 3.4 miles to a TDSR. Similar aggregated data for trucks 2 and 3 are also given in the table. Notice that while a DRT works from a single location, due to differences in debris type (and thus potentially different TDSR destinations for each load), the average distances by truck can vary.

Team *effectiveness* is then calculated as the total number of effective loads of debris hauled by a DRT on a given day, here $E = 8.3 + 12.7 + 9.1 = 30.1$ [from (2)]. Team *efficiency* is the average effective-load miles per truck: $\eta = (8.3 \times 3.4 + 12.7 \times 3.5 + 9.1 \times 4.7)/3 = 38.48$ [from (3)]. Finally, team *equality* is given by $\varepsilon = 1 - \sqrt{((8.3/10.03-1)^2 + (12.7/10.03-1)^2 + (9.1/10.03-1)^2)/(3\times 2) = 0.865}$ [from (5)].

As the effectiveness and efficiency measures are right skewed and cover a large range, the log (base 10) of each measure is used in the model. For equality, the range is on the unit interval and the data are skewed near 1. Thus, a square root transformation is applied to the inequality measure, and then the difference from 1 is taken. This transformation maintains the integrity of the equality measure because the square root is monotonic (thus preserving order) and the domain over the unit interval is still the unit interval. Finally, these transformed performance measures are then standardized to facilitate interpretation of the parameter estimates across dimensions.

IV. RESULTS

The main dataset for this study consists of detailed records from debris loads that were hauled by debris removal teams (DRTs) following the passage of an extensive tornado storm through the state of Alabama (USA) in April, 2011. Based on communications with USACE, the storm is estimated to have generated approximately 9.6 million cubic yards (7.3 million cubic meters) of debris. Data from a total of 132 300 load tickets issued from 9 May 2011 to 1 October 2011 were obtained from USACE in a machine-readable format, then cleaned and verified via automated and manual methods. For example, automated queries were used to identify potential anomalies and errors in the data (e.g., AM/PM time errors, date errors, out of range distances, invalid time sequences). Background materials on the event were collected and discussions with domain experts were held to provide additional contextual information on the dataset.

For the nearly five months of data considered here, the organization consists of one prime contractor and 33 subcontractors, operating within 33 municipalities (with a total of 49 TDSRs) in the state of Alabama. The number of active DRTs for the study period is given in the first panel of Fig. 2 (note that sudden dips occur on Sundays, when less work is done). An indication of fluidity in the system as a whole is shown in the second panel of Fig. 2, which depicts the percentage of trucks that switch DRTs each day. On average, 11% of the active trucks switch DRTs on any given day. An average of approximately 66 DRTs was active in the field on each day, and a total of 2271 different trucks took part in the mission. Team size ranged from two to 32 trucks (mean of 4.0), with 81.6% of teams consisting of five or fewer trucks.

Summary statistics for the untransformed study variables are given in Table II. As discussed previously, initial distributional testing on the study variables led to a log transformation being applied to the variables *Size*, *Effectiveness*, *Efficiency,* and *Equality* in order to satisfy normality assumptions of the linear modeling methods used for addressing each proposition.

When performing multiple inferences within a family of tests, a correction to the significance level used for each test is

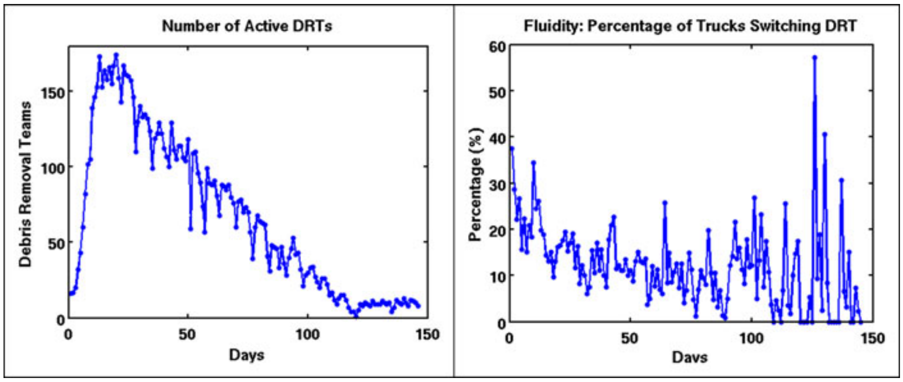

Fig. 2. Debris Removal Team Activity and Fluidity for Study Period.

TABLE II
SUMMARY OF UNTRANSFORMED KEY STUDY MEASURES

	Mean	*Std. Dev.*
Size	4.06	2.62
Fluidity	0.77	0.19
Effectiveness	11.28	10.42
Efficiency	18.12	13.33
Equality	0.86	0.11

required. Both tests (the first to assess Propositions 1 and 2, the second for Proposition 3) are considered of equal importance and thus the Bonferroni correction is used [27]. The propositions are assessed using two tests; thus, for a total $\alpha = 0.05$, the individual analyses are evaluated at $\alpha = 0.025$. In the statistical analysis that follows, the study sample size is 7583 team-days.

The empirical distributions for working history and size of DRTs are depicted in Fig. 3: panel 1 depicts a histogram of length of working history among all observed truck pairs, showing that the vast majority have working histories of two weeks or less; panel 2 shows that the number of trucks in a DRT rarely exceed ten.

To address Propositions 1 and 2, multivariate multiple regression is used [27], in which the response variable—team *performance*—is represented as a vector whose elements are the measures of *effectiveness*, *efficiency,* and *equality* described previously. In addition to the main predictor variables—team *fluidity* and team *size*—a single control variable, haul distance, is included, as payment per cubic yard differs depending on the distance between pickup and dropoff points. The current approach to measuring haul distance is consistent with the payment incentive structure, in that it assumes a straight line travel distance from pickup to drop-off point. It is not known how far this distance differs from actual travel distance; nonetheless, because distances are relatively small, deviation between actual and straight-line distance is not expected to be particularly great.

The results from the analysis for Propositions 1 and 2 are summarized in Table III. The table expresses parameter estimates in bold, notes p-values for statistically significant results, and provides standard errors of the parameter estimates in italics. Consistent with this analytic approach [27], the percentage of variation in the data explained by the model is reported along each dimension as *Adjusted* R^2.

The results show model significance across all dimensions of performance, although the percentage of variance explained across each dimension varies considerably: the predictor variables relate most closely to the *effectiveness* and *efficiency* dimensions (*Adj.*-R^2 of 0.460 and 0.403, respectively) relative to the *equality* dimensions (*Adj.*-$R^2 = 0.071$).

According to *Proposition 1.1*, increasing fluidity is expected to be associated with decreased effectiveness. This proposition is supported ($p = 0.0086$). Increasing fluidity is also associated with decreased efficiency ($p = 0.0090$) and decreased equality ($p = 0.0112$).

Proposition 1.2 states that, as the mission progresses, the effect of fluidity on the effectiveness dimension of performance will decrease. Therefore, in a comparison of the values of the coefficient for fluidity with respect to the effectiveness for period (t) versus period ($t + 1$), the proposition would be supported if the value for the latter period were greater than the value for the prior period (there are of course other approaches to this analysis). For completeness, the values of the fluidity coefficient of the model are provided in Fig. 4 along each of the three performance dimensions, as follows. The multivariate multiple regression model described previously was run on a recursive seven-day basis (i.e., data points separated by more than seven days have completely non-overlapping data) for all days, and the values of the coefficients plotted over time. The sample size varies across time starting at 321 teams, peaking at 1193 for the week starting at day 13. The sample size then slowly trails off to only 134 teams in the final week.

The data do suggest some temporal trends: there is a generally ascending tendency for effectiveness (thus supporting Proposition 1.2), as well as for efficiency (there is no suggested trend for equality). It should be emphasized, however, that all data series are perhaps best characterized as periodic. Looking more closely at the data, in roughly the first 40 days of the mission, the trends for effectiveness and efficiency are clearly upward, while the trend for equality is clearly downward. All three series track downward for roughly the period of days 40 to 60, after which time all track upward. The approximate midpoint of the mission may therefore be viewed as destabilizing for effectiveness and efficiency, with all trends upward following shortly after day 60. Thus, while the proposition is supported, future work may explore the reasons behind periodicity in the data, as well as those behind the destabilizing middle period of the series.

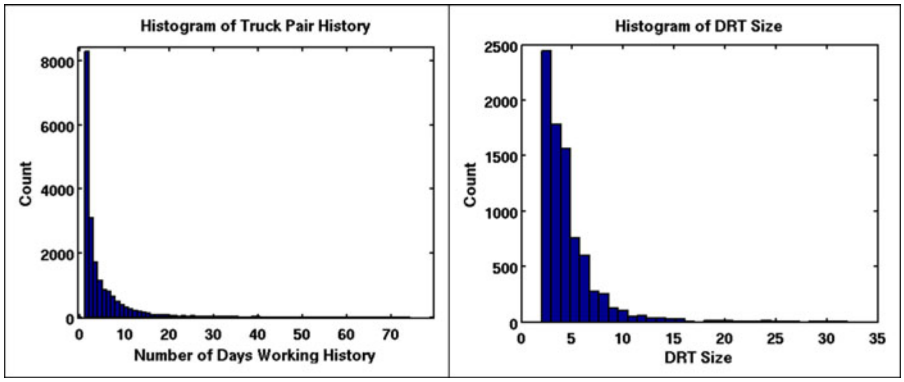

Fig. 3. Working History and DRT Size.

TABLE III
SUMMARY OF THE MULTIPLE MULTIVARIATE REGRESSION MODEL
(* INDICATES p < 0.001)

	Effectiveness	*Efficiency*	*Equality*
Fluidity	-0.062* (0.0086)	-0.082* (0.0090)	-0.191* (0.0112)
Size	0.665* (0.0086)	-0.0001 (0.99)	-0.138* (0.0112)
Haul Dist	-0.160* (0.0084)	0.631* (0.0088)	0.085* (0.0111)
Adj.- R^2	0.460	0.403	0.071

TABLE IV
SUMMARY OF EFFECTIVENESS VERSUS EFFICIENCY REGRESSION MODEL
(* INDICATES p < 0.001)

	Parameter Estimate
Fluidity	0.020* (0.0052)
Size	0.666* (0.0052)
Haul Dist	-0.790* (0.0051)
Adj. R^2	0.845

Next, according to *Proposition 2*, team size is expected to be positively associated with effectiveness and equality, but negatively associated with efficiency. As shown in Table III, consistent with expectation, increased team size is associated with increased effectiveness ($p = 0.0086$) but, contrary to expectation, also with reduced equality. Increased team size is not significantly associated with efficiency. Thus, Proposition 2 is partially supported.

Finally, *Proposition 3* states that increased team fluidity is expected to produce greater effectiveness than efficiency. Because all measures have been standardized, the difference between the effectiveness and efficiency standardized measures indicate a preference for effectiveness over efficiency. With an increase in fluidity, an increase in this difference would lead to support for the proposition. Results of a regression of the difference between effectiveness and efficiency on team composition are shown in Table IV. Consistent with the hypothesized relationship, fluidity has a significant positive effect ($p = 0.0052$) on this difference. Thus, Proposition 3 is supported.

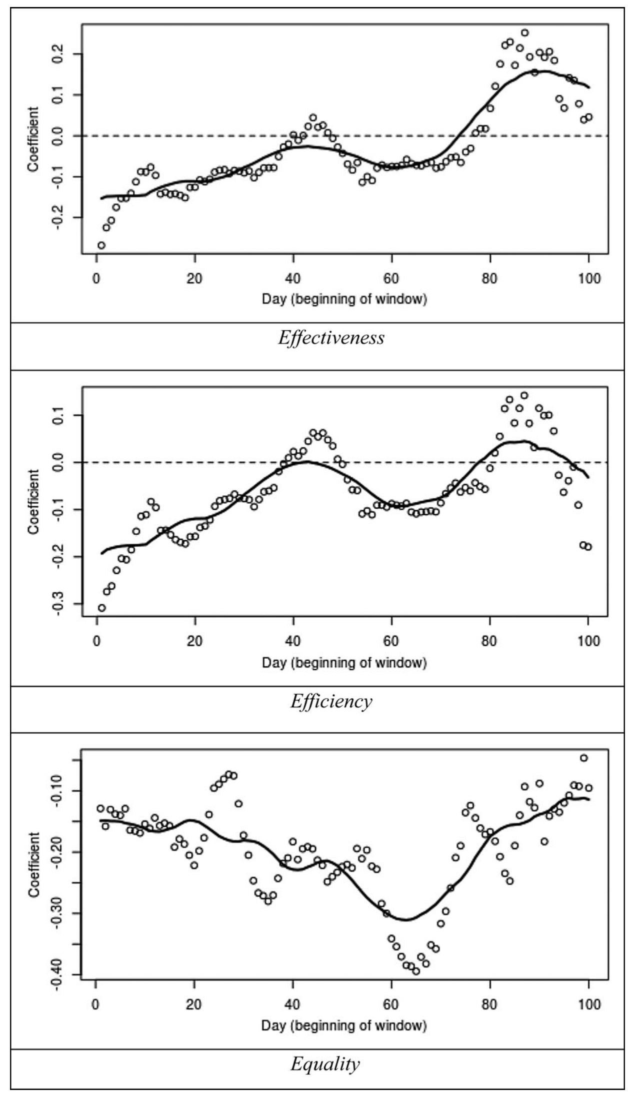

Fig. 4. Fluidity Coefficients for Performance Measures Over Recursive Week-Long Windows.

The Haul distance is significant along all dimensions of performance, with a negative effect on effectiveness, and a positive effect on both efficiency and equality.

V. Discussion

The results of the analysis of the effects of team composition on team performance show that increases in team size both contribute to and detract from performance, while increased fluidity consistently detracts from performance. From a broad perspective, then, these results suggest the potential explanatory power of tradeoff-based analysis in linking team composition to team performance. The high value of $Adjusted\text{-}R^2$ for the effectiveness and efficiency dimensions suggests that it is along these dimensions that fluidity and size are most relevant. However, given the potentially large number of transactions in systems of this kind, the significant (if small) effects along the equality dimension should also be considered.

Greater fluidity is associated with reduced performance along all three dimensions. It may be, for example, that higher turnover teams are subject to problems associated with learning of established routines, though exploration of the exact mechanisms that translate higher turnover into reduced performance is a matter for future work. Increased team size is associated with increased effectiveness, but reduced equality, suggesting a possible tradeoff between these two dimensions.

Graphical inspection of the relationship between fluidity and performance over time suggests a number of possible trends for each dimension, as well as the relationship among the dimensions themselves. Further research is needed to understand the nature of the relationship between size and the performance dimensions of effectiveness and equality in this context. For example, does increased team size lead to an increased need for coordination, thus taking resources away from task performance?

As suggested by the foregoing discussion, numerous opportunities exist for investigating team composition and performance using techniques derived from this study. The methods may be particularly well suited to the study of teams whose work processes are fairly stable, and where decisions on team composition are made exogenously. Possible examples include teams of construction workers, loggers, miners [17], and certain athletes [38], as well as teams of physicians in emergency rooms [2]. In contrast, these methods may not be appropriate for the study of so-called knowledge teams, where variety of work processes is endemic to operations, nor in settings where decisions on team composition are endogenous (i.e., by team members themselves). It should be emphasized that research along these lines depends on the availability of objective data on the execution of work processes by team members. Recent work on the development of advanced instrumentation for sporting contexts, for example, may help advance this research in domains such as tennis and football (soccer), particularly regarding the effects of fluidity.

Beyond questions of composition and performance are questions concerning the details of within-group processes (such as the factors that inform drivers' decisions about when, what, and how much debris to load at a pickup point). While the results of the present study explain a considerable degree of team performance, further work is needed in understanding patterns of cognition within teams in order to develop more robust models of performance.

As stated previously, considerable discussion surrounds definitions of teams versus other human collectives. The present study offers an example of how objective data collected during the execution of actual work processes may be used to ground some aspects of these discussions. This need is particularly acute as greater theoretical precision is brought to bear in developing the dimensions used to classify teams. Indeed, a limitation of the present research—one that is shared by many other studies of teams—is its reliance on a single definition of teamwork (albeit one that has achieved considerable traction). Additionally, it should be noted that while data such as these may support investigation along some aspects of, say, the team composition dimension (e.g., size), they may be less well suited to other aspects (e.g., attitudes/values).

Finally, as suggested by the aforementioned points, the results illustrate how teams may react to opportunity costs. Larger haul distances are associated with reduced effectiveness, but with improved efficiency and equality. This may reflect an overall intention by teams to extract the most value out of the resources available to them (here, the loading capacity of vehicles relative to travel time), while improving coverage over the area from which they are hauling debris.

VI. Conclusion

This study has explored performance effects associated with changes in team composition, focusing on potential tradeoffs along the dimensions of effectiveness, efficiency, and equality. Team composition, expressed in terms of team size and team fluidity, relates to all dimensions of performance in some way (with the exception of the effect of size on efficiency). A tradeoff between effectiveness and equality is suggested with respect to team size. Size and fluidity have differential effects on effectiveness, and a negative effect on equality. A possible mechanism underlying these results may relate to the nature of the work being performed. For production work, it may be more fruitful for team members to learn to adapt quickly to existing routines rather than innovating. Interestingly, the larger the team, the greater the reduction in effectiveness. Consequently, increasing team size as a means of offsetting the negative effects of turnover may actually accentuate these effects.

A second contribution of this study is in explicit modeling of performance tradeoffs in teamwork, thus extending previous research centered primarily on individual decision making. Additional domains—such as mining, logging, and athletics—may be amenable to investigation of some of these effects (particularly of fluidity), but, for reasons described previously, are likely to be restricted to settings where work processes are fairly stable and there are sufficient opportunities for instrumenting work processes.

Beyond these results, this study contributes to ongoing efforts to marshal detailed, process-level field data to address theoretically salient questions about teamwork processes. Possible extensions of this study are in modeling of system-level dynamics associated with adaptive behavior by teams. For example, one possible point of inquiry is in examining the mechanisms by which changes in size and fluidity are translated into variation in performance. Various information technologies, such as smart phones and geographic positioning systems, now enable such measurements to be taken accurately and inexpensively. Additional research with other instruments will likely be needed to explore cognitive-level effects within teams.

This research also contributes to the identification of opportunities for exploring models of control in the system under study. For example, currently, dispatchers allocate (and reallocate) individuals and equipment to teams in the field. There may be opportunities for using data on the results of these decisions to provide feedback to dispatchers in order to support subsequent allocation decisions. Moreover, dynamic analysis (as suggested previously) may be used to investigate the effect of these and other feedback effects on allocation decisions by dispatchers, as well as loading and routing and decisions by team members. These approaches will require explicit models both of the teams themselves and of the broader system within which they operate.

A third potential area of future work is in investigating team-level detection and response to disruptive events encountered in day-to-day operations. How a team addresses disruptions such as extreme weather and other hazardous or potentially hazardous conditions provides a measure of its flexibility, whether these disruptive events produce work stoppages, slowages, or simply replanning [19], [55]. In the context of the analysis of organizational performance, disruptive events offer a potential exogenous factor that may enable a deeper investigation of team performance.

References

[1] L. Argote and A. A. Kane, "Learning from direct and indirect experience in organizations," in *Group Creativity: Innovation Through Collaboration*, P. Paulus and B. Nijstad, Eds. New York, NY, USA: Oxford, 2003, pp. 277–303.

[2] C. M. Barnes, J. R. Hollenbeck, D. T. Wagner, D. S. DeRue, J. D. Nahrgang, and K. M. Schwind, "Harmful help: The costs of backing-up behavior in teams," *J. Appl. Psychol.*, vol. 93, no. 3, pp. 529–539, 2008.

[3] D. J. Beal, R. R. Cohen, M. J. Burke, and C. L. McLendon, "Cohesion and performance in groups: A meta-analytic clarification of construct relations," *J. Appl. Psychol.*, vol. 88, no. 6, pp. 989–1004, 2003.

[4] B. S. Bell and S. W. J. Kozlowski, "A typology of virtual teams: Implications for effective leadership," *Group Org. Manage.*, vol. 27, no. 1, pp. 14–49, 2002.

[5] H. Bernardin, C. Hagan, J. Kane, and P. Villanova, "Effective performance management: A focus on precision, customers, and situational constraints," in *Performance Appraisal: State of the Art in Practice*, J. W. Smither, Ed. San Francisco, CA, USA: Jossey-Bass, 1998, pp. 3–48.

[6] S. Bodily, "Police sector design incorporating preferences of interest groups for equality and efficiency," *Manage. Sci.*, vol. 24, no. 12, pp. 1301–1313, 1978.

[7] E. K. Browning and W. R. Johnson, "The trade-off between equality and efficiency," *J. Political Economy*, vol. 92, no. 2, pp. 175–203, 1984.

[8] E. Carmel and R. Agarwal, "Tactical approaches for alleviating distance in global software development," *IEEE Softw.*, vol. 18, no. 2, pp. 22–29, Mar./Apr. 2001.

[9] M. R. Carrell and J. E. Dittrich, "Equity theory: The recent literature, methodological considerations, and new directions," *Academy Manage. Rev.*, vol. 3, no. 2, pp. 202–210, 1978.

[10] P. B. Coulter, "Measuring the inequity of urban police services: A methodological discussion with applications," *Policy Studies J.*, vol. 8, no. 5, pp. 683–698, 1980.

[11] L. A. DeChurch and J. R. Mesmer-Magnus, "The cognitive underpinnings of effective teamwork: A meta-analysis," *J. Appl. Psychol.*, vol. 95, pp. 32–53, 2010.

[12] T. L. Dickinson and R. M. McIntyre, "A conceptual framework of teamwork measurement," in *Team Performance Assessment and Measurement: Theory, Methods, and Applications*, M. T. Brannick, S. E , and C. Prince, Eds. Mahwah, NJ: Lawrence Erlbaum Associates, 1997, pp. 331–356.

[13] J. A. Espinosa, S. A. Slaughter, R. E. Kraut, and J. D. Herbsleb, "Familiarity, complexity, and team performance in geographically distributed software development," *Org. Sci.*, vol. 18, no. 4, pp. 613–630, 2007.

[14] B. Golany and E. Tamir, "Evaluating data envelopment approach analysis," *Manage. Sci.*, vol. 41, no. 7, pp. 1172–1184, 1995.

[15] B. Golany and E. Tamir, "Evaluating efficiency-effectiveness-equality trade-offs: A data envelopment analysis approach," *Manage. Sci.*, vol. 41, no. 7, pp. 1172–1184, 1995.

[16] P. S. Goodman, E. Ravlin, and M. Schminke, "Understanding groups in organizations," in *Research in Organizational Behavior*, L. L. Cummings and B. M. Staw, Eds. Greenwich, CT, USA: JAI Press, 1987, pp. 121–173.

[17] P. S. Goodman and S. Shah, "Familiarity and work group outcomes," in *Group Process and Productivity*, S. Worchel, W. Wood, and J. A. Simpson, Eds. Newbury Park, CA, USA: Sage, 1992, pp. 276–298.

[18] M. Grabowski and K. H. Roberts, "Risk mitigation in virtual organizations," *Org. Sci.*, vol. 10, no. 6, pp. 704–721, 1999.

[19] G. Grote, J. C. Weichbrodt, H. Günter, E. Zala-Mezö, and B. Künzle, "Coordination in high-risk organizations: The need for flexible routines," *Cognition, Technol. Work*, vol. 11, no. 1, pp. 17–27, 2009.

[20] D. H. Gruenfeld, P. V. Martorana, and E. T. Fan, "What do groups learn from their worldliest members? Direct and indirect influence in dynamic teams," *Organizational Behavior Human Decision Processes*, vol. 82, no. 1, pp. 45–59, 2000.

[21] J. W. Guthrie, Jr, H. A. Priest, Jr, and E. Salas, "The continued evolution of team research: A theoretical model of performance in multiteam systems," in *Proc. Human Factors Ergonom., Soc. Annu. Meeting*, 2005, pp. 582–585.

[22] R. A. Guzzo and M. W. Dickson, "Teams in organizations: Recent research on performance and effectiveness," *Annu. Rev. Psychol.*, vol. 47, no. 1, pp. 307–338, 1996.

[23] J. R. Hackman, *Leading Teams: Setting the Stage for Great Performances*. Boston, MA, USA: Harvard Business School Press, 2002.

[24] D. A. Harrison, K. H. Price, and M. P. Bell, "Beyond relational demography: Time and the effects of surface- and deep-level diversity on work group cohesion," *Acad. Manage. J.*, vol. 41, no. 1, pp. 96–107, 1998.

[25] J. R. Harrison, Z. Lin, G. R. Carroll, and K. M. Carley, "Simulation modeling in organizational and management research," *Acad. Manage. Rev.*, vol. 32, no. 4, pp. 1229–1245, 2007.

[26] R. D. Huseman, J. D. Hatfield, and E. W. Miles, "A new perspective on equity theory: The equity sensitivity construct," *Acad. Manage. Rev.*, vol. 12, no. 2, pp. 222–234, 1987.

[27] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, 3rd ed. Upper Saddle River, NY, USA: Prentice-Hall, 1992.

[28] S. W. J. Kozlowski and B. S. Bell, "Work groups and teams in organizations," in *Handbook of Psychology (Vol. 12): Industrial and Organizational Psychology*, W. C. Borman, D. R. Ilgen, and R. J. Klimoski, Eds. New York, NY, USA: Wiley-Blackwell, 2001, pp. 333–375.

[29] B. Latane, K. Williams, and S. Harkins, "Many hands make light the work: The causes and consequences of social loafing," *J. Personality Social Psychol.*, vol. 37, no. 6, pp. 822–832, 1979.

[30] J. M. Levine, R. M. Moreland, L. Argote, and K. M. Carley, "Personnel turnover and team performance," US Army Research Institute for the Behavioral and Social Sciences, 2005.

[31] L. G. Luther, *Disaster Debris Removal after Hurricane Katrina: Status and Associated Issues*. Washington, DC, USA: Congressional Research Service, Library of Congress, 2006.

[32] M. B. Mandell, "Modelling effectiveness-equity trade-offs in public service," *Manage. Sci.*, vol. 37, no. 4, pp. 467–482, 1991.

[33] M. T. Marsh and D. A. Schilling, "Equity measurement in facility location analysis: A review and framework," *Eur. J. Oper. Res.*, vol. 74, no. 1, pp. 1–17, 1994.

[34] J. Mathieu, M. T. Maynard, T. Rapp, and L. Gilson, "Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future," *J. Manage.*, vol. 34, no. 3, pp. 410–476, 2008.

[35] J. E. Mathieu and G. Chen, "The etiology of the multilevel paradigm in management research," *J. Manage.*, vol. 37, no. 2, pp. 610–641, 2011.
[36] J. E. Mathieu, M. A. Marks, and S. J. Zaccaro, "Multi-team systems," in *International Handbook of Work and Organizational Psychology*, N. Anderson, D. Ones, S. H. K. , and C. Viswesvaran, Eds. London, U.K.: Sage, 2001, pp. 289–313.
[37] M. Mohtashami, S. Cheickna, I. Il, and F. P. Deek, "A study of information processing effectiveness in inter-organisational collaboration in software development," *Int. J. Inform. Technol. Manage.*, vol. 10, no. 2, pp. 124–150, 2011.
[38] A. A. Morgan-Lopez, L.A. Cluff, and W. Fals-Stewart, "Capturing the impact of membership turnover in small groups via latent class growth analysis: Modeling the rise of the New York Knicks of the 1960 s and 1970 s," *Group Dynamics: Theory, Res. Practice*, vol. 13, no. 2, pp. 120–132, 2009.
[39] National Research Council, Facing Hazards and Disasters: Understanding Human Dimensions. Washington, DC, USA: National Academy Press, 2006.
[40] J. R. Rizzo, R. J. House, and S. I. Lirtzman, "Role conflict and ambiguity in complex organizations," *Administ. Sci. Quart.*, vol. 15, no. 2, pp. 150–163, 1970.
[41] E. Salas, C. S. Burke, and K. C. Stagl, "Developing teams and team leaders: Strategies and principles," in *Leader Development for Transforming Organizations*, D. Day, S. F. Zaccaro, and S. M. Halpin, Eds. Mahwah, NJ, USA: Lawrence Erlbaum Associates, 2004.
[42] E. Salas, N. J. Cooke, and M. A. Rosen, "On teams, teamwork, and team performance: Discoveries and developments," *Human Factors: J. Human Factors Ergonom. Soc.*, vol. 50, no. 3, pp. 540–547, 2008.
[43] E. Salas, T. L. Dickinson, S. A. Converse, and S. I. Tannenbaum, "Toward an understanding of team performance and training," in *Teams: Their Training and Performance*, R. W. Sweezey and E. Salas, Eds. Norwood, NJ, USA: Ablex, 1992, pp. 3–29.
[44] E. Salas, K. C. Stagl, and C. S. Burke, "25 years of team effectiveness in organizations: Research themes and emerging needs," *Int. Rev. Ind. Organizational Psychol.*, vol. 19, pp. 47–91, 2004.
[45] C. S. Saunders and M. K. Ahuja, "Are all distributed teams the same? Differentiating between temporary and ongoing distributed teams," *Small Group Res.*, vol. 37, no. 6, pp. 662–700, 2006.
[46] G. Schreyögg and J. Sydow, "Organizing for fluidity? Dilemmas of new organizational forms," *Org. Sci.*, vol. 21, no. 6, pp. 1251–1262, 2010.
[47] I. Steiner, *Group Processes and Productivity*. San Diego, CA, USA: Academic, 1972.
[48] K. G. Stoneman and A. M. Dickinson, "Individual performance as a function of group contingencies and group size," *J. Organizational Behavior Manage.*, vol. 10, no. 1, pp. 131–150, 1989.
[49] E. Sundstrom, M. McLntyre, T. Halfhill, and H. Richards, "Work groups: From the Hawthorne studies to work teams of the 1990 s and beyond," *Group Dynamics: Theory, Res., Practice*, vol. 4, no. 1, pp. 44–67, 2000.
[50] A. M. Townsend, S. M. DeMarie, and A. R. Hendrickson, "Virtual teams: Technology and the workplace of the future," *Academy Manage. Executive*, vol. 12, no. 3, pp. 17–29, 1998.
[51] U. S. Congress, *A Failure of Initiative: Final Report of the Select Bipartisan Committee to Investigate the Preparation for and Response to Hurricane Katrina*. Washington, DC, USA: United States House of Representatives, 2006.
[52] R. Wagemen, "How leaders forster self-managing team effectiveness: Design choices versus hands-on coaching," *Org. Sci.*, vol. 12, no. 5, pp. 559–577, 2001.
[53] K. E. Weick, "Systematic observational methods," in *Handbook of Social Psychology*, G. Lindzey and E. Aronson, Eds. New York, NY, USA: Random House, 1985, pp. 567–634.
[54] J. L. Wildman, A. L. Thayer, M. A. Rosen, E. Salas, J. E. Mathieu, and S. R. Rayne, "Task types and team-level attributes: Synthesis of team classification literature," *Human Resource Develop. Rev.*, vol. 11, no. 1, pp. 97–129, 2012.
[55] D. Woods, "Essential characteristics of resilience," in *Resilience Engineering: Concepts and Precepts*, E. Hollnagel, D. Woods, and N. Leveson, Eds. Aldershot, U.K.: Ashagate, 2006.

David Mendonça (M'01–SM'12) received the B.A. degree from the University of Massachusetts at Amherst, Amherst, MA, USA, in 1990, the M.S. degree in public policy and management from Carnegie Mellon University, Pittsburgh, PA, USA, in 1994, and the Ph.D. degree in decision sciences and engineering systems from Rensselaer Polytechnic Institute, Troy, NY, USA in 2001,

He is currently an Associate Professor in the Department of Industrial and Systems Engineering, Rensselaer Polytechnic Institute. He holds a secondary appointment in the Department of Cognitive Science, Rensselaer Polytechnic Institute. He has also been a Visiting Scholar at the Delft University of Technology (The Netherlands) and the University of Lisbon (Portugal). His main research interests include the cognitive foundation of individual and group decision making in high stakes, time-pressured conditions, particularly during postdisaster emergency response.

James D. Brooks (M'02) received the B.S. degree in electrical and computer engineering from Grove City College, Grove City, PA, USA, in 2004, and the M.S. degree in electrical and computer engineering from the Georgia Institute of Technology, Atlanta, GA, in 2009. He is currently working toward the Ph.D. degree from the Department of Industrial and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA.

Since 2010, he has been a member of the Real Time Optimization and Controls Laboratory, General Electric Global Research further developing rail automation, building energy management systems, and naval vessel power management systems. Previously, he was a Systems Engineer at General Electric Transportation from 2004 to 2010. His research interests include methods for distributed control of complex human–machine systems (particularly those having queueing dynamics) and human–automation interaction.

Martha Grabowski received the B.S. degree in marine transportation/nautical science from the U.S. Merchant Marine Academy, Great Neck, NY, USA, in 1979, the Ph.D. degree in management/information systems, the M.B.A. degree, and the M.S. degree in industrial engineering from Rensselaer Polytechnic Institute, Troy, NY, USA, in 1987 1982, and 1983, respectively,

She is the McDevitt Chair in information systems, and a Professor and the Director of the Information Systems Program in the Madden School of Business, Le Moyne College in Syracuse, NY. She is also a Research Professor in the Department of Industrial and Systems Engineering, Rensselaer Polytechnic Institute in Troy, NY. A member of the American Bureau of Shipping. His research focuses on the impact of technology on individuals, groups and organizations in complex, safety-critical systems, risk analysis and risk mitigation in large-scale systems, and the role of human and organizational error in high consequence settings.

Dr. Grabowski is a Licensed Former Merchant Officer and retired LCDR in the US Naval Reserve.

She is a Lifetime National Associate of the National Academies/National Research Council, and is currently chairing the National Academies' policy study assessing the U.S.' Arctic Oil Spill Response capabilities for eight sponsoring agencies.