Journal of Applied Psychology
2013, Vol. 98, No. 3, 454–468

0021-9010/13/$12.00 DOI: 10.1037/a0031909

Voluntary Survey Completion Among Team Members: Implications of Noncompliance and Missing Data for Multilevel Research

Robert R. Hirschfeld
University of Colorado at Colorado Springs

Michael S. Cole
Texas Christian University

Jeremy B. Bernerth and Tracey E. Rizzuto
Louisiana State University

We explored whether voluntary survey completion by team members (in aggregate) is predictable from team members' collective evaluations of team-emergent states. In doing so, we reanalyze less-than-complete survey data on 110 teams from a published field study, using so-called traditional and modern missing data techniques to probe the sensitivity of these team-level relationships to data missingness. The multivariate findings revealed that a greater within-team participation rate was indeed related to a higher team-level (mean) score on team mental efficacy (across all four missing-data techniques) and less dispersion among team member judgments about internal cohesion (when the 2 modern methods were used). In addition, results show that a commonly used approach of retaining only those teams with high participation rates produces inflated standardized effect size (i.e., R^2) estimates and decreased statistical power. Suggestions include research design considerations and a comprehensive methodology to account for team member data missingness.

Keywords: composition models, deletion, multiple imputation, response rate, systematic nonresponse parameter

Among multilevel researchers, it is standard practice to collect data (survey responses) from individual team members and then use an appropriate composition model to create higher (team-) level constructs (Chan, 1998, 2011; Cole, Bedeian, Hirschfeld, & Vogel, 2011). When conducting such aggregation, a majority of researchers interested in group or team phenomena have adopted a consensus composition model, wherein the average of scores among team members' survey responses is used to operationalize the team-level constructs (Chan, 1998). Nevertheless, evidence also supports dispersion among team members' judgments as a theoretically and empirically meaningful team-level construct (Cole et al., 2011; Lindell & Brandt, 2000). Consequently, the use of a dispersion-composition model (Chan, 1998, 2011) for capturing variation among team members' responses is gaining recognition as a viable alternative when examining team-level variables derived from data originating from individual team members (Meade & Eby, 2007).

Although consensus and dispersion composition are the fundamental models for data aggregation in multilevel (team) research, an important limitation of both forms of composition is their reliance on survey responses from teams' individual members. Nesterkin and Ganster (2012), for example, have suggested that it is nearly impossible to obtain a within-team participation rate of 100% across all teams in a field research sample. The ensuing problem is that when only some team members complete a survey, questions emerge about whether the obtained data adequately represent the opinions of the entire team (Maloney, Johnson, & Zellmer-Bruhn, 2010). That is, less-than-complete team member representation may be insufficient as a basis for summarizing team phenomena. Indeed, prior results suggest that when mean-based variables (consensus composition) are derived from incomplete data, and the missing team member data are missing not at random, the consequences are biased (i.e., understated) estimates and equivocal findings (Maloney et al., 2010; Timmerman, 2005). Furthermore, in those instances when variation-based variables (dispersion composition) are derived from incomplete data, and the data are missing for any reason (i.e., either completely at random or not at random), biased (i.e., understated) estimates and distorted findings will be the result (Allen, Stanley, Williams, & Ross, 2007a, 2007b; Newman & Sin, 2009). Hence, team members' survey (non)response behavior (i.e., within-team participation rate) is an important consideration for researchers interested in collective phenomena that arise from bottom-up or emergent processes (Newman, 2009).

This article was published Online First March 4, 2013.

Robert R. Hirschfeld, Department of Management, College of Business, University of Colorado at Colorado Springs; Michael S. Cole, Department of Management, Texas Christian University; Jeremy B. Bernerth, Rucks Department of Management, Louisiana State University; Tracey E. Rizzuto, School of Human Resource Education and Workforce Development, Louisiana State University.

The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, the Department of Defense, or the United States government. The authors appreciate suggestions and guidance offered by Arthur G. Bedeian, Christopher M. Berry, Richard P. DeShon, and Frederick P. Morgeson.

Correspondence concerning this article should be addressed to Robert R. Hirschfeld, College of Business, University of Colorado at Colorado Springs, 1420 Austin Bluffs Parkway, Colorado Springs, CO 80918. E-mail: rrhirschfeld@gmail.com

The application of consensus- and dispersion-composition models is further complicated by the absence of clear decision guidelines for handling team member missingness or nonparticipation (Newman, 2009). Accordingly, it is not surprising that Maloney et al.'s (2010) review of 62 team-based articles revealed "wide variation in how researchers handled within-group nonresponse" (p. 287). Nevertheless, a relatively common practice for handling missing data is the application of a team-retention rule based on within-team participation rates (Stanley, Allen, Williams, & Ross, 2011). This approach applies an arbitrary retention rule whereby any team that does not achieve a minimum survey participation rate is subsequently "dropped" or "filtered out" of the sample. Pertinent evidence for the prevalence of this practice has been provided by Allen et al. (2007a). They found, for example, that when a study's researchers provided information about how less-than-complete teams were handled, most used a retention rule to exclude teams on the basis of a *minimum* percentage of respondents (ranging from at least 40% to 100% representation).

Although the removal of low-representation teams is a frequent practice, a number of scholars (e.g., Biemann, Cole, & Voelpel, 2012; Bliese & Halverson, 1998; Newman, 2009) have repeatedly advocated that team-level analyses be conducted with all available data rather than data from only high-representation teams. This paradox suggests that the many calls for multilevel research to incorporate advanced methodological techniques, stemming from the now extensive literature on missing survey data (Allison, 2009; Newman, 2009; Stanley et al., 2011), have been largely unheeded. Given that aggregation of data from an incomplete set of team members can produce inaccurate statistical results, it follows that a principal objective of the present study was to explore, from a team-level perspective, the nature and implications of within-team research participation.

In pursuing this objective, our study contributes to the literature in three ways. First, it directly responds to calls for research exploring whether, and to what extent, within-team participation rate (which varies among teams) is associated with emergent team-level (viz., mean-based and dispersion-based) phenomena. Timmerman (2005) first raised the issue that team member missingness (i.e., nonparticipation by a proportion of team members) may reveal something meaningful about team dynamics, and yet acknowledged that this possibility has not received empirical attention. Likewise, Newman and Sin (2009, p. 138) have observed that the identification of any associations between team-level survey participation rates and the bottom-up constructs being researched would represent a critical step forward in advancing multilevel research. Recognizing the theoretical and practical value of understanding why research participation rates differ among teams, we echo Newman's (2009) sentiment that it is important to ultimately catalog the extent to which within-team participation rate is connected to team-level phenomena.

A second contribution is that we explore the consequences resulting from the common practice of retaining only those teams with a relatively high within-team participation rate. Specifically, we explore whether "filtering out" low-representation teams alters the ability of a set of team-level predictors to explain the important outcome of teamwork effectiveness. Drawing on empirical evidence that the costs of excluding low-representation teams will far outweigh this practice's assumed benefits, we expect our findings to highlight the often misunderstood impact of these conventional team-retention rules on a study's conclusions.

A third contribution is that we illustrate how two contemporary missing data techniques can be effectively used for probing whether team composition-model variables derived from an incomplete representation of a team's members are biasing a study's substantive findings (i.e., sensitivity analyses). The first contemporary method is multiple imputation (Sinharay, Stern, & Russell, 2001) of data that are missing on a construct-by-construct or measure-specific basis (i.e., some relevant data are available for all participants). A second contemporary method entails quantitatively adjusting team-level means and standard deviations using nonlocal systematic nonresponse parameters (Newman, 2009; Rogelberg, Luong, Sederberg, & Cristol, 2000). Notably, this second method can be used even when there are no data available on the missing persons (i.e., survey-level missingness). In essence, we demonstrate a comprehensive approach that enabled us to more effectively use all available data and thereby empirically rule out the possibility that team-level findings were biased by missing individual-level (i.e., team member) data. By doing so, we hope that the proposed techniques will likewise shape future research applications by offering practical procedures for dealing more effectively with missing team member survey data.

Theory and Propositions

Insight into the nature of team-based missingness promises to be an important complement to what is already known about the potential consequences of individualized missingness (Nesterkin & Ganster, 2012; Stanley et al., 2011). Although scholars have speculated that there are likely some team-level factors that explain differences among teams in their within-team participation rates (e.g., Newman, 2009), we could not find any published field research on the extent to which team members' survey (non)response behavior is linked to team-level phenomena. Rogelberg and Stanton (2007) have suggested, however, that if certain teams have lower response rates than other sampled teams, this "may be indicative of some important underlying differences . . ." (p. 204). In keeping with this idea, Newman (2009) has suggested that volitional noncompletion of a research survey by potential participants may occur for substantive reasons involving group dynamics, to include collective social-psychological factors. Similarly, Maloney et al. (2010) offered a hypothetical example in suggesting that members of teams with better interpersonal interactions would more likely complete a survey about their teams.

For the present purposes, then, a first step is to identify substantive team-level constructs (emergent phenomena) that plausibly relate to team member survey completion in predictable ways (Maloney et al., 2010; Rogelberg & Stanton, 2007). According to London and London (2007), some of the more important representations of team functioning are team confidence and cohesion. This reasoning provides the basis for the present study's focus on team efficacy and team internal cohesion. *Team efficacy* is conceptualized as a team's belief that it can perform well on a specified type of task (Stajkovic, Lee, & Nyberg, 2009). Complementing team efficacy, *internal cohesion* represents a general collective spirit or sense of unified striving among members of a team (Levi, 2007; Rosh, Offermann, & Van Diest, 2012). Mathieu, Maynard, Rapp, and Gilson (2008) classified both of these constructs as team-emergent states because they represent shared or emergent phenomena (bottom-up processes) that are common to all members of a team. A distinct attribute of emergent states is that they materialize at the team level of analysis via mutual interdependence and coordinated interaction (i.e., from discontinuous compilations of various inputs; Kozlowski & Klein, 2000), and owing to their unique motivational properties they can explain team outcomes in ways not accounted for by other types of team variables (Mathieu et al., 2008). As such, the emergent states of team efficacy and internal cohesion play a central role in the present research endeavor.

We draw from two theoretical perspectives in positing connections between the emergent states of teams (efficacy and cohesion) and voluntary research participation by team members. The perspectives differ in the nature of the collective experience thought to motivate individual team members' participation in a research study, and, thus, each suggests unique ways in which team-emergent states may be linked to within-team participation rates. First, the *social psychology of prosocial behavior* suggests that the rate of within-team research participation is related positively to our focal emergent team states. In exploring this perspective, we use mean-based aggregation—representing consensus composition—to operationalize the team-level variables that index efficacy and cohesion. A second theoretical perspective draws on principles of *normative influence* to suggest that within-team research participation reflects group forces that both homogenize perceptions and encourage behavioral uniformity. In examining this latter perspective, we use the variability among team members' collective judgments—representing dispersion composition—to operationalize the team-level variables of efficacy and cohesion.

Mean of Team Members' Perceptions of the Team: Proposition 1

To provide a grounded team-level explanation of why members of some teams will be more likely to comply with an organization-based request to complete a survey about their team, we first draw on the theory of prosocial behavior. The logic underpinning the social psychology of prosocial behavior suggests that a high rate of within-team participation would be an additive result of *independent cooperation* (or compliance; Rogelberg et al., 2000) by individual team members. Completing a research survey is a cooperative act that entails a small personal sacrifice for the benefit of others or an external entity (see, e.g., Dovidio, Piliavin, Schroeder, & Penner, 2006). The proclivity of team members to display such cooperation or compliance may stem from positive interpersonal relationships experienced within the team (Reich & Hershcovis, 2011). According to Flynn (2011), for example, when individuals experience support from others, an ensuing sense of goodwill motivates them to generally respond in kind. To the degree that interpersonal support is experienced routinely within a team, as reflected in high mean levels of confidence and cohesion, individual members are likely to hold favorable perceptions of their social environment and thereby have a positive outlook toward interactions with others. This valued sense of group belongingness and benevolence should thus inspire individual members to cooperate with others both within and outside the team (see De Cremer & van Knippenberg, 2002; Reich & Hershcovis, 2011). Accordingly, we propose that teams experiencing relatively high mean levels of these shared social-psychological states (viz., team efficacy and cohesion) will have members who more likely complete voluntary research surveys.

> *Proposition 1:* The proportion of team members voluntarily participating in a team survey study will be greater for teams with higher mean levels of team efficacy and/or cohesion.

Diversity Among Team Members' Perceptions of the Team: Proposition 2

It has long been known that group processes can shape homogeneous cognition and behavior among members of a collective (Zajonc, 1965). Therefore, a second way to predict survey response behavior within a team is from the notion that *normative influence* engenders more uniform (or less diverse) perceptions among a team's members. This perspective is supported by the well-established literature on the implications of team context for the behavior of team members (Hackman, 1992). According to Levy and Nail (1993), for example, normative influence shapes and reinforces *echo behavior* in social environments. As such, it seems reasonable to theorize that the social and perceptual factors associated with this form of social contagion should promote the adoption of predominant attitudes and the imitation of normative behavior by members of a team. With a shared baseline expectation about their team (i.e., a less disperse outlook), team members would converge in their interpretations of and responses to a given situation (De Jong & Dirks, 2012). In this vein, we anticipate that descriptive social norms (Goldstein & Cialdini, 2011) encouraging benevolence would engender greater agreement by team members to participate in a survey study (Groves, Singer, & Corning, 2000; Newman, 2009). The implication of such normative influence is that less diverse judgments among team members (a more uniform mindset) about our focal team-emergent states should be associated with a higher rate of within-team research participation.

> *Proposition 2:* The proportion of team members voluntarily participating in a team survey study will be greater for teams with less diverse team member perceptions about their team's efficacy and/or cohesion.

Implications of Sample Representation for Research on Teams: Proposition 3

As previously noted, the most common approach to handling missing data is the application of a team-retention rule (Stanley et al., 2011). A team-retention rule is based on within-team research participation rates, in which a minimum percentage of respondents is required within each team for it to be retained. In using such a rule, researchers recognize that team member missingness (survey nonresponse) can introduce measurement error into team-level variables derived from composition models. Therefore, a higher within-team participation rate can be simulated by using only those teams with a relatively high proportion of representation. According to Stanley et al., the underlying assumption for team-retention rules is that "it is better to have no information from a group than to have information that is contaminated by high levels of measurement error" (p. 510).

Nevertheless, methodologists (e.g., Newman, 2009) have criticized the practice of "dropping" or "filtering out" teams with a relatively low level of research participation. The most obvious cost associated with this approach is that it deletes teams from the sample, leading to loss of statistical power (Chan, 2011). A second, less obvious, issue is that it can artificially homogenize the team-level variables of interest (Cole et al., 2011), thereby producing estimates of team-level relationships that depart from the true relationships (Nesterkin & Ganster, 2012; Stanley et al., 2011). Consequently, when a study on team phenomena uses a team-retention rule on the basis of within-team response rates, it potentially compromises the accuracy of standardized effect size estimates (cf. Preacher, Rucker, MacCallum, & Nicewander, 2005).

Proposition 3: Excluding teams with a relatively low rate of within-team research participation from a sample will (a) alter how much criterion (viz., observed teamwork effectiveness) variance is explained by the predictor set and (b) diminish the statistical significance of predictor parameters as a result of less statistical power.

A Primer on the Original Research Context

We used a data set previously used by Hirschfeld and Bernerth (2008). These published data are ideal for various reasons. First, the field data represent a relatively large sample (N = 110) of action teams encompassing trainees in a 5-week team-centered U.S. Air Force officer development program (ODP). Data from such a setting is valuable for the present study in that, according to Nesterkin and Ganster (2012), only one published study to date (see Timmerman, 2005) has used field data to explore the nature and implications of team member nonresponse in research examining team constructs emerging from bottom-up processes. Another advantageous feature is that ODP teams were synchronously formed in that the teams' creation, development, and dissolution occurred in tandem. Furthermore, by design, the teams were relatively equivalent in composition from the standpoint of demographic characteristics, job classifications, and military career status. These unique study attributes are important because they effectively control for potential confounds, including teams' temporal stage of development (Marks, Mathieu, & Zaccaro, 2001) and team faultlines (Thatcher & Patel, 2011). The 110 teams ranged in size from 12 to 15 members, with 85 of the teams (77%) having 14 members.

Research Design and Data Collection Procedures

Other favorable features of the data stem from the original research design. In particular, a two-wave lagged survey design was used when gauging team members' responses; two team-emergent states (team mental efficacy and team physical efficacy) were assessed in Week 2 using Survey 1, and a third emergent state (internal cohesion) was measured about 3 weeks later using Survey 2. On each survey occasion, all of the ODP trainees were seated in a large auditorium and invited to voluntarily complete a paper-and-pencil survey for the benefit of external research on the ODP. Given this survey design, a comparison of survey data results derived from team member listwise (survey data are included from only those *individuals* providing complete survey data across both surveys) relative to pairwise (whatever survey data are available from those *individuals* are included) deletion techniques is possible.

Although Hirschfeld and Bernerth's (2008) primary research question was considered relevant enough to be approved by the pertinent Air Force human subjects research review board, it was not deemed mission critical to the U.S. Air Force. As such, survey administration complied with the human subjects ethical principle of *respect for persons* (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1978), meaning that potential participants were free to make their own choice about partaking in the study. Nevertheless, all ODP trainees were required to be in attendance on both survey administration dates. On each occasion, information about the research and its intended benefits were communicated, and potential participants were asked to place their surveys (completed or not) into one of several boxes inside the auditorium. The purpose of this arrangement was to ensure that no one in attendance would immediately realize whether an individual team member completed his or her survey. Therefore, the ODP trainees were presented with a survey condition free of coercion or undue influence, and no personal incentive or benefit was offered for completing surveys.

As part of the surveying process, potential participants were asked to provide their ODP identification numbers, which provided a means of pairing participants' responses on Surveys 1 and 2. This surveying approach uses what are known as *identified employee surveys* (Saari & Scherbaum, 2011). Notably, the ODP trainees were asked to provide their identification numbers but were not required to do so. A number of experts prefer this approach to administering identified employee surveys (e.g., Black, Hyland, & Rutigliano, 2011; Froelich, 2011; Saari & Scherbaum, 2011). Official records indicated the precise number of individuals composing each team (by identification number), thereby allowing us to determine every team's survey participation rate and whether each team member was compliant or noncompliant.

Method

Team-Level Predictors

Mean-based variables. Survey 1 (Week 2) contained the measures of *team mental efficacy* and *team physical efficacy*, each of which consisted of seven items with a 6-point response format ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). Survey 2 (Week 5) contained a measure of *internal cohesion*, consisting of six items with a response scale ranging from 1 (*strongly disagree*) to 6 (*strongly agree*). The items composing these measures delineated the team as an explicit referent (i.e., a referent-shift consensus model; Chan, 1998). Using all available survey data, coefficient alpha was .88 for mental efficacy, .91 for physical efficacy, and .90 for internal cohesion.

We examined the statistical adequacy of aggregating individual members' responses to the team level by testing whether average scores differed significantly across teams, as indicated by one-way analyses of variance (ANOVAs) and by calculating intermember reliability (ICC1 and ICC2). The ANOVAs, using team as the independent factor, demonstrated that there is between-team variance for all three variables ($p < .01$). Moreover, the values of ICC1 and ICC2 were .35 and .86 for team mental efficacy, .25 and .79 for team physical efficacy, and .40 and .88 for internal cohesion. Coupled with ANOVA results, ICC estimates demonstrate that the focal mean-based predictors represent emergent team-level properties that are distinct from their individual-level parent constructs (see LeBreton & Senter, 2008, for a detailed discussion of ICC estimates). We therefore created team-level mean scores using the predominant, and simplest, aggregation method that entails computing a weighted arithmetic mean for each team by averaging across the team member responses.[1]

Dispersion-based variables. The within-team standard deviation (SD_{WG}) is a preferred index of within-group dispersion (Newman & Sin, 2009; Roberson, Sturman, & Simons, 2007). Accordingly, we computed dispersion scores for teams' mental efficacy, physical efficacy, and internal cohesion by using the square root of the variance for each set of team scores. Consequently, higher scores represent greater dispersion or diversity in perceptual ratings among team members. A brief analysis revealed that variations in teams' size did not affect the inferences drawn from our analyses (see Cole et al., 2011).

Team-Level Criteria

Team participation rate. Within-team participation rate is the criterion for our tests of Proposition 1 and Proposition 2. We operationalized team participation rate as the proportion of team members who fully participated in the research endeavor.[2] For our study's purposes, the measure of participation represents the crux of cooperation as described by Dovidio et al. (2006). That is, cooperating with the research requests entailed a small personal sacrifice of (a) devoting one's time to complete a survey on each of two occasions, (b) making one's judgments known to the researchers by way of specific responses to survey items, and (c) writing one's correct ODP identification number on each of the two surveys. The average within-team participation rate was 73% (min = 36%; max = 100%; median = 77%).

Observed teamwork effectiveness. Teams were assigned an independent observer for the duration of the ODP. At the conclusion of the 5-week ODP, those observers evaluated the effectiveness of their team's teamwork (see Hirschfeld & Bernerth, 2008, for a detailed description). We used this assessment of observed teamwork effectiveness as the criterion for examining Proposition 3. This outcome variable serves an important role in that it was derived from an external source, and it was available for all 110 teams (i.e., no missing data).

As described in the original study, behavioral examples of extremely low and high levels were provided for each of four aspects of teamwork effectiveness, which observers used in rating their team from *extremely low* (1) to *extremely high* (6). The four aspects were (a) level of effort toward task accomplishment, (b) level of commitment to help one another so that team performance is maximized, (c) degree to which team members communicated effectively, and (d) prevalence with which team members used resources (e.g., time, member expertise) well during team tasks. The four ratings were averaged to compute an overall score; coefficient alpha was .87, and the global score ranged from 1.50 to 6.00 ($M = 4.17$; $SD = .95$) across the 110 teams.

Noncompletion of Surveys and the Issue of Team Representation

To investigate our study's propositions, it was necessary to have some teams with less than perfect (100%) within-team participation rates. In the present instance, the noncompletion of surveys by a team's individual members was most likely an active rather than a passive form of nonparticipation (or nonresponse; Rogelberg et al., 2003). That is, by virtue of the captive audience with time scheduled to complete both surveys distributed to each member of every team (attendance was mandatory for all ODP officer participants), a nonrespondent made a deliberate choice to not cooperate with a request to complete the survey that was administered. Hence, it is quite possible that missing survey data at the team level resulted from purposeful noncompliance by one (or more) of a team's members (Rogelberg et al., 2000).

Given these conditions, we attempted to index the nature and magnitude of data missingness by estimating a systematic nonresponse parameter (SNP) known as d_{miss} (calculated as the standardized respondents–nonrespondents mean difference on a variable; Newman & Sin, 2009). In this regard, Rogelberg et al. (2003) have outlined four approaches for estimating the degree to which respondents providing complete data differ systematically from nonrespondents; these approaches are the archival, follow-up, wave, and intentions techniques (see Table 1 of their article, on p. 1105). In regard to these four approaches, we were able to locally estimate two individual-level SNPs by way of the *archival approach* because Hirschfeld and Bernerth's (2008) original database contained personal identification numbers. More specifically, we

[1] We recognize that this approach to aggregation neither considers the accuracy of individual team members' judgments nor includes such a consideration in the aggregation process. van Bruggen, Lilien, and Kacker (2002) have asserted that although the simple averaging of individual members' responses is "the most common practice in empirical organizational research," aggregation procedures that assign greater weight to those participants who seemingly provide more accurate survey responses "significantly improves the accuracy of organizational response data" (p. 470). On the basis of van Bruggen et al. (2002), we conducted supplemental analyses using an alternative aggregation procedure that may be referred to as a response data-based weighted mean. We defined and operationalized each team member's judgment inaccuracy by its deviation from the team's simple arithmetic mean (see Equation 2). Then, we assigned a weight to the individual members' judgments such that greater emphasis was placed on judgments with smaller deviations (see Equation 3). Finally, we computed weighted team-level means for all teams (see Equation 4). Results of empirical tests using these response data-based weighted means were virtually identical to those we obtained when using the simple arithmetic means for each team. Nevertheless, it should be noted that an anonymous reviewer suggested the use of the van Bruggen et al. equation could produce some currently unknown consequences. As such, we report results involving only the simple arithmetic mean-based variables.

[2] When the criterion variable is a proportion, it is important to screen the data so as to determine an appropriate regression method. For example, the relationship between a continuous predictor and a proportion response variable may be sigmoidal (i.e., a flattened S) rather than linear. When a sigmoidal curve is present in the data, it is customary to apply an alternative regression model (e.g., probit regression or beta regression). In the present instance, however, an examination shows that the data on within-team participation rate fall primarily between .3 and .8, otherwise known as the middle, or *linear* section of the sigmoidal curve. As such, given that in our case the proportion response variable was effectively linear, we used ordinary least squares regression. We deemed this to be the most straightforward approach to testing our propositions.

created an archival database on all of the ODP trainees that contained performance scores on an exam of teamwork concepts (multiple-choice format) and ratings on observed leadership potential (as evaluated by an external observer). We then compared respondents (as a group) with nonrespondents (as a group) on each of these two indicators (Rogelberg et al., 2003). Such comparisons are beneficial in that the magnitude of any observed differences helps to inform the missing data scenarios used when probing whether a study's findings are somehow biased as a result of data missingness. Cohen's d effect size values for these comparisons were $d_{miss} = -.27$ for performance on the exam and $d_{miss} = -.33$ for leadership potential. In each case, noncompliant individuals achieved lower scores as compared with what the respondents had attained (the average d_{miss} was $-.30$). This evidence that d_{miss} is negative is important, in that it mirrors prior conclusions that active nonrespondents differ substantively (e.g., in terms of negative attitudes and pessimistic outlook) from respondents (Newman, 2009; Rogelberg et al., 2003, 2000). As explained in more detail below, estimates of d_{miss} can be taken into account when modern missing data techniques are used.

Data Analytic Issues When Survey Data Are Missing[3]

As in much of applied psychological research, missing data are common in quantitative research on groups or teams (Maloney et al., 2010). With the term *missing data*, we are referring to data that are missing for some but not all of a team's members on one or more composition-model variables. As we noted previously, within-team participation needed to vary among sampled teams so that some teams would have a rate of research participation falling below 100%. Nevertheless, the accuracy of sample statistics and ensuing hypothesis tests may be undermined when data are missing for team-level variables (viz., mean-based and dispersion-based) derived from composition models (Allen et al., 2007a; Stanley et al., 2011). To comprehensively address these issues, in testing our three propositions we used so-called traditional as well as modern techniques for handling missing data (see Allison, 2009, and Schafer & Graham, 2002, for detailed reviews). We thus followed the important recommendation that multiple methods be used in a single study to determine the sensitivity of substantive findings to data missingness (e.g., Allison, 2009).

For the present investigation, any missing survey data are missing on a construct-by-construct basis (see Newman, 2009). This means that although one or more of a team member's responses are missing for at least one survey measure, some data on each person are still available. Recall that the original study entailed two surveys administered 3 weeks apart. Whereas Survey 1 (Week 2) contained the measures of *team mental efficacy* and *team physical efficacy*, Survey 2 (Week 5) contained a measure of *internal cohesion*. Thus, survey data can be missing for a team member on Survey 1, Survey 2, or both survey waves. As we previously explained, however, archival data (e.g., knowledge test scores and leadership ratings) were available for all individuals regardless of their survey participation.

Traditional missing data techniques. As conventional methods, listwise deletion and pairwise deletion of data from individuals are the two most widely applied missing data approaches (Peugh & Enders, 2004). *Listwise deletion* is the blanket removal of individual cases (i.e., team members) with any missing data on the variables of interest. Accordingly, listwise deletion of data from some team members is carried out before the remaining data from individuals are aggregated to the team level of analysis. The rationale for listwise deletion is that one should aggregate survey data from only those individual team members who were by all accounts earnest in their efforts to provide complete and useful data on each study variable. A potential downside of listwise deletion is an increased likelihood of measurement error accompanying fewer individual member responses per team (Maloney et al., 2010). Thus, a commonly applied solution to increase team member representation is to use so-called pairwise deletion of missing data. *Pairwise deletion* refers to the use of all available data from individuals; any discarding of cases (i.e., team members) takes place on a construct-by-construct basis (before data from individuals are aggregated to the team level of analysis) so that team-level properties reflect as many of the team's individual members' responses as possible. We use both of these conventional missing data approaches in our exploration.

Modern missing data techniques. We used two additional techniques for the purpose of sensitivity analyses. *Modern Method 1* addresses any missing data through *m*ultiple *i*mputation (*MI*). A number of empirical studies have demonstrated the superiority of *MI* over the traditional techniques of listwise and pairwise deletion (Newman, 2009; Sinharay et al., 2001). In fact, *MI* has been described as a "state of the art" (Schafer & Graham, 2002, p. 147) technique for analyzing data sets with measure-specific missing data. Formally, *MI* is a process that produces m imputed data sets, each of which includes "filled in" values based on a random draw from a distribution of probable missing values. Of direct relevance to the present investigation, *MI* has been explicitly recommended for dealing with instances of data missingness within teams (e.g., Newman & Sin, 2009, p. 137).

To conduct *MI*, we used all of the survey data available. The average within-team nonresponse observed for Survey 1 (16%) and Survey 2 (18%) is similar to the rate of noncompliance within previously published individual-level research (15%, Rogelberg et al., 2003; 14%, Spitzmüller, Glenn, Sutton, Barr, & Rogelberg, 2007). Then, we used the individual-level data to create $m = 5$ imputed data sets using the Markov Chain Monte Carlo (MCMC) algorithm. Allison (2009) has noted that MCMC is the most widely applied *MI* method and has also suggested that five imputed data sets are sufficient to obtain efficient estimates (see also Rubin, 1987). To create the five imputed data sets, we used the SPSS Missing Values Analysis software package. In regards to the imputation model, we entered a superset of 11 variables encompassing the three emergent states (viz., mental efficacy, physical efficacy, and internal cohesion) and eight auxiliary variables (team members' age, gender, source of entry into the U.S. Air Force, military career status, teamwork knowledge test score, and observed leadership potential score, as well as their team's size and gender composition).

As a result of these steps, we had $m = 5$ full individual-level data sets; that is, team members' scores on mental efficacy, phys-

[3] The ensuing text and subsequent analyses pertaining to issues of missing data at the team level were primarily guided by comments from anonymous reviewers and the associate editor. The authors are grateful for the guidance received.

ical efficacy, and internal cohesion now existed for any individuals who did not respond to a given survey measure (i.e., imputed data) as well as for those who did complete the measure (i.e., observed data). Each full individual-level data set was then aggregated to the team level of analysis using standard procedures, thereby creating *mean-based* composition-model variables as well as *dispersion-based* composition-model variables. With the imputed data sets aggregated to the team level, we performed all substantive analyses on each of the $m = 5$ (full) team-level data sets and then averaged the resulting parameter estimates across m analyses to form a final set of estimates (Rubin, 1987). Finally, we used Rubin's equation (as applied by Allison, 2002, see Equation 5.1, p. 30) to compute improved standard errors (Newman, 2003, p. 340) and corrected t values for the regression coefficients.

Modern Method 2 does not use *MI* methods, but rather quantitatively adjusts the obtained team-level scores by using *s*ystematic *n*onresponse *p*arameters (i.e., *team-level* QA_{SNP}). Our use of SNPs (d_{miss}) for composition-model variables allowed us to effectively "manage" the potential systemic bias that can arise from survey noncompliance. To elaborate, important insight into survey noncompliance by individuals has been provided by several exemplary studies conducted by Rogelberg et al. (2000; 2003). These studies show, for example, that active nonrespondents make a conscious decision to not comply with a survey request, and, as a result, this class of nonrespondents has been found to hold an outlook that is negative relative to that of passive nonrespondents (e.g., individuals who forgot to complete the survey) as well as respondents (see Rogelberg et al., 2003, 2000; Spitzmüller, Glenn, Barr, Rogelberg, & Daniel, 2006; Spitzmüller et al., 2007). This active nonresponse is considered a form of nonrandom missingness (i.e., an individual's missingness is related to one's standing on study variables), which creates nonresponse bias and undermines the generalizability of findings. This is why Newman and Sin (2009) assert that data missing not at random is the "worst form of missingness" (p. 115).

The treatment of any observed missing data as necessarily reflecting the active form of nonparticipation should yield robust and conservative tests of our study's propositions, because applying SNPs (d_{miss}) to our composition-model variables is a way to account for systematic missingness (Newman, 2009). As we reported earlier, the archival database suggested that noncompliants' measured performance was lower (average $d_{miss} = -.30$) as compared with the ODP trainees who provided complete survey data. In addition, research conducted by Rogelberg et al. (2003, 2000) establishes estimates of d_{miss} ranging from $-.30$ to $-.60$ (see Newman & Sin, 2009, pp. 115–116). We applied both of these empirical estimates as nonlocal d_{miss} values, which enabled us to further probe how sensitive results were to varying (moderate to extreme) levels of d_{miss} (Newman & Sin, 2009) and, thus, ascertain how sensitive our results are to missing data even at the somewhat extreme d_{miss} value of $-.60$.

We therefore adjusted the observed team-level mean scores for the three emergent states at two levels of d_{miss} (see Equation 8 in Newman & Sin, 2009, p. 121). As an example, a team's mean score on team mental efficacy has two vectors in the data set—one adjusted at $d_{miss} = -.30$ and the other adjusted at $d_{miss} = -.60$. We likewise adjusted the observed SD_{WG} for the three team-emergent states. In doing so, we applied Newman and Sin's nonlinear transformation (see Equation 6, p. 118) to calculate $SD_{WG\ complete}$. We again used values of $d_{miss} = -.30$ and $d_{miss} = -.60$ in two separate equations to obtain adjusted estimates of SD_{WG}.

Results

Table 1 presents the means, standard deviations, and intercorrelations for the study variables derived from pairwise deletion of survey data (i.e., all available survey data were included). The potential controls (i.e., Variables 1–7) listed in Table 1 are described fully in Hirschfeld and Bernerth (2008). The intercorrelations in Table 1 reveal a relative absence of relationships between

Table 1
Descriptive Statistics and Correlations Based on all Available (Pairwise) Data

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Team-level controls																
1. Number of team members	13.73	0.59	—													
2. Proportion of team that is female	0.17	0.05	−.23	—												
3. Team knowledge pool	86.45	2.44	−.05	−.23	—											
4. Project X Phase 1 Results	4.21	1.56	.06	−.15	.26	—										
5. Project X Phase 2 Results	4.67	1.36	.14	−.04	.08	.37	—									
6. Problem-solving results	10.21	5.20	.07	.00	.24	−.04	.05	—								
7. Field operations results	15.20	6.25	.05	−.19	.16	.05	.15	−.06	—							
Mean-based predictors																
8. Team mental efficacy	4.78	0.48	.33	−.19	.16	.59	.26	.24	.10	—						
9. Team physical efficacy	4.68	0.43	.26	−.24	.03	.45	.17	.03	.20	.71	—					
10. Internal cohesion	4.80	0.55	−.01	−.20	.21	.31	.22	.14	.30	.46	.43	—				
Dispersion-based predictors																
11. SD_{WG} mental efficacy	0.57	0.18	−.20	.15	−.18	−.32	−.18	−.19	.09	−.56	−.28	−.26	—			
12. SD_{WG} physical efficacy	0.66	0.18	−.27	.25	−.01	−.29	−.13	−.13	.03	−.44	−.43	−.21	.55	—		
13. SD_{WG} internal cohesion	0.62	0.20	−.07	.07	−.24	−.19	−.19	−.19	−.16	−.26	−.11	−.60	.31	.20	—	
Team-level criteria																
14. Within-team participation rate	0.73	0.16	−.01	.07	.18	.18	.09	.18	.00	.24	.05	.24	−.10	−.03	−.26	—
15. Observed teamwork effectiveness	4.17	0.95	.08	−.10	.45	.30	.26	.26	.37	.46	.29	.53	−.30	−.21	−.42	.27

Note. $N = 110$ teams. Pairwise survey data were used, which encompass 84% of possible survey data for the mental efficacy measures and 82% for internal cohesion. The mean-based predictors were derived from a simple arithmetic method, and the dispersion-based predictors were unadjusted. Correlations with an absolute value of .19 or greater are significant at $p < .05$. SD_{WG} = within-team standard deviation.

the potential control variables and within-team participation rate. Consequently, we excluded these factors from analyses on Propositions 1 and 2, as Becker (2005) has shown that including unnecessary control variables may yield biased parameter estimates.[4] Yet, the intercorrelations do reveal significant relationships of several control variables with observed teamwork effectiveness. Therefore, the potential controls are included in the analyses we use for exploring Proposition 3, as explained more fully below.

Propositions 1 and 2: Systematic Team-Level Relationships

Given that the mean-based and dispersion-based components of the same team-level construct are statistically dependent (Chan, 2011; Cole et al., 2011), we explored Proposition 1 and Proposition 2 by entering all focal predictors into the same series of regression equations. Table 2 presents results from the use of listwise deletion and pairwise deletion of survey data within teams. Table 3 shows results from the use of two modern methods of accounting for missing data.

Proposition 1 suggests that the mean-based components of the teams' emergent states will be related positively to within-team participation rates. Results presented in Tables 2 and 3 show that team mental efficacy was related positively to within-team participation across all techniques for handling missing data (βs ranging from .33 to .43, $p < .05$). Although team physical efficacy was related negatively to within-team participation rate (an instance of statistical suppression; Cohen & Cohen, 1983, p. 94) when traditional techniques for missing data were used (see Table 2), it was not when more advanced missing data methods were used (see Table 3). Finally, internal cohesion was not related to team participation rate in any equation.

Proposition 2, which maintains that within-team dispersion on the same three emergent states will be related negatively to within-team participation rate, received mixed support across the missing data methods. To elaborate, results based on traditional techniques for handling missing data do not support Proposition 2; that is, none of the dispersion-based variables (SD_{WG}) significantly predicted within-team participation rate (see Table 2). In contrast, when the two modern methods were used to account for team member missingness (see Table 3), internal-cohesion dispersion ($SD_{WG\ complete}$) was a significant predictor (and in the anticipated direction) of within-team participation rate when *MI* was used ($\beta = -.33$, $p < .01$) and when *team-level* QA_{SNP} with a d_{miss} of $-.60$ was used ($\beta = -.28$, $p < .05$). Statistical significance was not attained, at a level of .05, when *team-level* QA_{SNP} with a d_{miss} of $-.30$ was used ($\beta = -.20$, $p = .10$). Across the three modern-method equations, the simple average of the t values for internal-cohesion dispersion ($SD_{WG\ complete}$) is -2.25, which equates to $p < .05$.

Table 2
Predicting Within-Team Participation Rate: Hirschfeld and Bernerth's (2008) Survey Data

Variable	Listwise survey data R^2	β	t	Pairwise survey data R^2	β	t
Model	.13*			.14*		
Team mental efficacy		.40	2.81**		.43	2.70**
Team physical efficacy		−.31	−2.22*		−.27	−1.87†
Internal cohesion		.15	1.23		.10	.78
SD_{WG} mental efficacy		.11	.84		.13	1.02
SD_{WG} physical efficacy		−.07	−.53		.02	.16
SD_{WG} internal cohesion		−.11	−.99		−.16	−1.29

Note. $N = 110$ teams. Listwise survey data encompass 73% of possible survey data. Pairwise survey data encompass 84% of possible survey data for the mental efficacy measures; 82% for internal cohesion. The mean-based predictors were derived from a simple arithmetic method, and the dispersion-based predictors were unadjusted. SD_{WG} = within-team standard deviation.
† $p < .10$. * $p < .05$. ** $p < .01$.

Proposition 3: Implications for the Findings of a Team-Level Explanatory Model

Turning to *Proposition 3*, it states that excluding teams with a relatively low rate of within-team research participation from a sample will (a) alter how much criterion (viz., observed teamwork effectiveness) variance is explained by the predictor set and (b) diminish the statistical significance of predictor parameters as a result of less statistical power. To examine this proposition, and for purposes of comparison, we estimated a series of regression equations that mirror the model predicting observed teamwork effectiveness in Table 4 of Hirschfeld and Bernerth (2008, p. 1434). This model encompasses the seven control variables described previously, and teams' mental and physical efficacy along with internal cohesion. This 10-predictor model of teamwork effectiveness seems ideal to use in examining Proposition 3, because the criterion is an important construct, measured by an external source with no missing data, and a considerable amount of criterion variance is explained (i.e., 50% R^2 is a robust basis for comparison).

To test Proposition 3, we examined standardized effect size (R^2) and the statistical significance of predictor parameters (β) for the sample of 110 teams and compared them with the corresponding estimates for a subsample of high-representation teams. We first estimated a series of regression equations to establish baseline results on the entire sample of 110 teams. For the next step of generating comparative results from a subsample of high-representation teams, we applied a team-retention rule defined as a within-team participation rate of greater than 75%.[5] By doing this, we "dropped" or "filtered out" 53 teams from the sample (i.e., roughly 50% of the total sample of teams). Using this subsample of 57 high-representation teams, we estimated a second set of regression equations using the same 10-predictor model. Table 4

[4] Results with control variables essentially mirror the reported results and are available from the authors upon request.

[5] Our use of a 75% team-retention rule reflects the dual aim of maintaining statistical power for the high-representation subsample (which will encompass fewer teams) while also minimizing the overlap between the subsample of 57 teams and the complete sample of 110 teams (see Stanley et al., 2011, for a detailed discussion). This dual aim is accomplished by using a cutoff value (75%) that is close to the average within-team participation rate for the complete sample of teams (73%).

Table 3
Predicting Within-Team Participation Rate: Two Modern Methods of Accounting for Missing Survey Data

Variable	Multiple imputation			QA_{SNP} (d_{miss} of −.30)			QA_{SNP} (d_{miss} of −.60)		
	R^2	β	t	R^2	β	t	R^2	β	t
Model	.19**			.16**			.20***		
Team mental efficacy		.34	2.15*		.40	2.48*		.33	2.08*
Team physical efficacy		−.23	−1.47		−.23	−1.58		−.17	−1.18
Internal cohesion		−.03	−.24		.09	.72		.06	.48
$SD_{WG\ \mathrm{complete}}$ mental efficacy		.01	.11		.11	.87		.07	.55
$SD_{WG\ \mathrm{complete}}$ physical efficacy		−.06	−.16		−.01	−.07		−.05	−.43
$SD_{WG\ \mathrm{complete}}$ internal cohesion		−.33	−2.84**		−.20	−1.64		−.28	−2.27*

Note. $N = 110$ teams. Multiple imputation of missing team member survey data represents Modern Method 1 for taking missing data into account. The two QA_{SNP} procedures (each of which incorporates a different estimate of d_{miss}) are conducted at the team level, and together they represent Modern Method 2; the obtained team-level survey data scores are quantitatively adjusted (*QA*) by using systematic nonresponse parameters (*SNPs*) in the form of d_{miss}. SD_{WG} = within-team standard deviation.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

presents results across the missing data techniques we used for the present study.[6]

With the sample of 110 teams, and collectively considering the array of missing data techniques used, the set of 10 predictors accounted for an average of 52% of the variance in teamwork effectiveness. For comparison, also included in Table 4 are the regression results from only the 57 teams with at least a 75% within-team participation rate. Using data from this subsample of teams, the set of 10 predictors accounted for an average of 56% of the variance in teamwork effectiveness. Put differently, for the high-representation subsample of 57 teams, the explained criterion variance was 8% greater (.56/.52) than what it was for the 110 team sample. Furthermore, for the entire sample of 110 teams, the focal variables of team mental efficacy (*ps* $< .05$) and internal cohesion (*ps* $\leq .01$), as well as the control variables of team knowledge pool (*ps* $\leq .001$) and field operations (*ps* $\leq .01$), were all significant predictors of teamwork effectiveness. For the high-representation subsample of teams, however, team knowledge pool was the only predictor achieving statistical significance at $p < .05$. Collectively, these results provide strong support for Proposition 3.

Discussion

This study is timely in that the frequency with which team-based research has been published continues to surge (Hollenbeck, Beersma, & Schouten, 2012). As not enough is known about the nature and implications of missingness in multilevel research involving teams (Nesterkin & Ganster, 2012), we empirically explored potential associations between team-level attributes (i.e., mean- and dispersion-based emergent states) and within-team survey participation rate. The findings that ensued are the first to demonstrate linkages between teams' emergent states (bottom-up processes) and voluntary research participation among a team's members. As such, our study supports the notion that team members' decisions to (not) complete organizational surveys may represent a reaction to the quality of their team-based experiences (e.g., social-psychological processes, group dynamics). Furthermore, our findings highlight the importance of estimating multiple missing data models to ascertain how sensitive study results are to model choice. This underscored theme is consistent with the extensive literature on individual-level missing data (Allison, 2009; Newman, 2009), and it has important theoretical and practical implications that apply to any research involving composition-model constructs.

Implications for Theory on Team Phenomena and Research Participation

We sought to pursue a better understanding of why members of some teams are more (less) likely than those of other teams to voluntarily complete surveys. Our results support the relevance of theoretical and empirical findings from the team-emergent states literature (e.g., Mathieu et al., 2008) for the study of team members' survey response behavior. Mean-based team mental efficacy, an important emergent state known to promote teams' problem solving and performance, predicted the proportion of team members who completed surveys. This finding was consistent across all four methods used to account for missing survey data, and it supports the notion that within-team survey participation is a form of generalized compliance or cooperation (Dovidio et al., 2006). As such, it seems that when a team's members experience positive interpersonal interactions within a given setting (e.g., as a team member within the ODP), the motivated goodwill that ensues may extend beyond the immediate team to include prosocial conduct toward others who appear within that same, or a similar, setting (see Flynn, 2011). In a future study, it might be interesting to explore whether individuals' research cooperation versus noncompliance relates more strongly to prosocial behavior directed toward

[6] At the request of an anonymous reviewer, we tested whether the observed results for the sample of 110 teams versus those for the high-representation subsample of 57 teams could be explained by lower reliability of the group means (i.e., ICC2) in the 110 team sample. This possibility arises because the entire sample had a median of two fewer respondents per team, compared with what existed for the high-representation subsample of 57 teams. Recall that for the sample of 110 teams, the values of ICC2 were .86 for team mental efficacy, .79 for team physical efficacy, and .88 for internal cohesion. In comparison, for the high-representation subsample of 57 teams, values of ICC2 were .85 for team mental efficacy, .83 for team physical efficacy, and .91 for internal cohesion. Given that .02 is the mean absolute difference on ICC2 (across the three variables) between the sample and subsample, it is unlikely that the minimal difference unduly influences our findings for Proposition 3.

Table 4
Predicting Observed Teamwork Effectiveness

	Listwise survey data			Pairwise survey data		
Sample of 110 teams	R^2	β	*t*	R^2	β	*t*
Model	.50			.52		
Number of team members		.04	.45		.03	.32
Proportion of team that is female		.11	1.36		.11	1.47
Team knowledge pool		.30	3.73***		.31	3.87***
Project X Phase 1 Results		.01	.14		−.02	−.24
Project X Phase 2 Results		.07	.87		.07	.94
Problem-solving results		.10	1.20		.08	.98
Field operations results		.25	3.27**		.24	3.17**
Team mental efficacy		.28	2.19*		.32	2.50*
Team physical efficacy		−.12	−1.08		−.11	−1.05
Internal cohesion		.29	3.24**		.30	3.49***

	Listwise survey data			Pairwise survey data		
Subsample of 57 teams	R^2	β	*t*	R^2	β	*t*
Model	.56			.56		
Number of team members		−.06	−.48		−.07	−.65
Proportion of team that is female		.04	.32		.05	.40
Team knowledge pool		.30	2.59*		.31	2.69**
Project X Phase 1 Results		.11	.76		.08	.59
Project X Phase 2 Results		.09	.84		.09	.80
Problem-solving results		.21	1.75†		.21	1.73†
Field operations results		.17	1.54		.17	1.60
Team mental efficacy		.25	1.20		.27	1.31
Team physical efficacy		−.14	−.82		−.11	−.69
Internal cohesion		.27	1.96†		.25	1.91†

	Multiple imputation			QA_{SNP} (d_{miss} of −.30)		
Sample of 110 teams	R^2	β	*t*	R^2	β	*t*
Model	.53			.52		
Number of team members		.01	.10		.03	.33
Proportion of team that is female		.11	1.50		.11	1.44
Team knowledge pool		.30	3.77***		.30	3.81***
Project X Phase 1 Results		−.04	−.39		−.03	−.26
Project X Phase 2 Results		.07	.88		.07	.94
Problem-solving results		.07	.91		.08	.97
Field operations results		.24	3.19**		.24	3.15**
Team mental efficacy		.36	2.70**		.31	2.44*
Team physical efficacy		−.14	−1.27		−.10	−.96
Internal cohesion		.30	3.23**		.30	3.52***

	Multiple imputation			QA_{SNP} (d_{miss} of −.30)		
Subsample of 57 teams	R^2	β	*t*	R^2	β	*t*
Model	.57			.56		
Number of team members		−.09	−.77		−.08	−.68
Proportion of team that is female		.05	.40		.05	.40
Team knowledge pool		.31	2.72**		.31	2.68**
Project X Phase 1 Results		.06	.44		.08	.58
Project X Phase 2 Results		.09	1.28		.09	.79
Problem-solving results		.20	1.71†		.21	1.72†
Field operations results		.18	1.63		.18	1.61
Team mental efficacy		.31	1.48		.28	1.33
Team physical efficacy		−.12	−.78		−.11	−.68
Internal cohesion		.25	1.87†		.25	1.93†

(*table continues*)

Table 4 (*continued*)

Sample of 110 teams	QA_{SNP} (d_{miss} of −.60) R^2	β	t
Model	.52		
Number of team members		.03	.33
Proportion of team that is female		.11	1.39
Team knowledge pool		.30	3.75***
Project X Phase 1 Results		−.03	−.26
Project X Phase 2 Results		.07	.93
Problem-solving results		.08	.97
Field operations results		.24	3.12**
Team mental efficacy		.30	2.35*
Team physical efficacy		−.09	−.86
Internal cohesion		.31	3.55***

Subsample of 57 teams	QA_{SNP} (d_{miss} of −.60) R^2	β	t
Model	.56		
Number of team members		−.08	−.71
Proportion of team that is female		.05	.40
Team knowledge pool		.31	2.67**
Project X Phase 1 Results		.08	.56
Project X Phase 2 Results		.09	.78
Problem-solving results		.21	1.72
Field operations results		.18	1.62
Team mental efficacy		.28	1.36
Team physical efficacy		−.11	−.68
Internal cohesion		.26	1.94†

Note. All regression models encompass the same predictors and criterion. Across the same missing data techniques, results for the complete sample of 110 teams can be compared to the results for the high-representation subsample of only 57 teams. For the first seven predictor variables, global scores were derived from objective team-level indicators (see Hirschfeld & Bernerth, 2008). The final three predictor variables are mean-based composition-model variables measured by way of surveys. Listwise survey data encompass 73% of possible survey data for all three focal predictors. Pairwise survey data encompass 84% of possible survey data for team mental efficacy and team physical efficacy, and 82% for internal cohesion. *QA* = quantitatively adjusted; *SNPs* = systematic nonresponse paramenters in the form of d_{miss}.
† $p \leq .10$. * $p \leq .05$. ** $p \leq .01$. *** $p \leq .001$.

a team's members relative to that toward other entities such as the organization as a whole.

Our investigation of dispersion-based predictors lends mixed support for the notion that survey completion among a team's members can be examined using the theoretical framework of normative influence. When the two modern methods were used to account for missing survey data, internal-cohesion dispersion ($SD_{WG\ complete}$) was a significant predictor of within-team participation rate across two of the three regression equations. A temporal consideration may help explain why dispersion about internal cohesion (measured in Week 5), and not team mental or physical efficacy (measure in Week 2), was the basis of the tentative support for Proposition 2. Specifically, research has previously shown time-lagged effects for attitudinal convergence among team members and its role in team behavior (Rizzuto, Mohammed, & Vance, 2011). Furthermore, because team members' coordinated interactions denote the elemental content (i.e., raw material; Kozlowski & Klein, 2000) of bottom-up emergence, any phenomena arising from normative influence are probably more likely to be revealed over time. For this reason, longitudinal research may be particularly useful in teasing out differential relationships between dispersion-based predictors and research participation among a team's members.

Finally, an important caveat when interpreting our study's findings is that we did not directly investigate specific cognitive and emotional factors (within individuals) that might have played a role in determining whether individuals decided to participate, or not, in the survey study. As such, there remains much to learn regarding specifically what intrapersonal factors may impel research participation. For the present study, the reasons for not complying with requests to complete research surveys may have differed among the noncompliant team members. For instance, whereas some noncompliant individuals may have been concerned about confidentiality, others may have simply been tired. By including an array of relevant explanatory factors, future research could provide unique insight into why only some team members are willing to provide data for a study. We thus encourage researchers to directly explore reasons for survey noncompliance, by means of methods described by Rogelberg and colleagues (2003).

Implications for Multilevel Research Methods

Given the absence of evidence-based team-retention guidelines (Maloney et al., 2010; Newman & Sin, 2009), many researchers have applied percentage-based retention rules to govern their decisions about retaining or dropping teams from a database. As has been discussed here and elsewhere (e.g., Stanley et al., 2011), the practice of "filtering out" teams with a relatively low participation rate essentially creates a quasi-artificial sample consisting of teams that are homogenous on the construct(s) of interest. Our findings reveal that this commonly applied practice can have inadvertent ramifications for model specification, standardized effect size, and the interpretability of results. As anticipated, results affirm that excluding teams on the basis of a minimum percentage of respondents per team results in diminished statistical power and distorted criterion variance (model R^2) explained. This practice is not unlike a problematic sampling procedure referred to as the "extreme groups approach," which involves eliminating data (i.e., those in the middle of a sample distribution) and is known to inflate standardized effect size estimates (Preacher et al., 2005). In considering the findings of our study as well as those from prior research (e.g., Maloney et al., 2010; Nesterkin & Ganster, 2012), we conclude that limiting a research sample to high-representation teams is inappropriate. That is, deleting teams from a sample goes against "a simple yet fundamental principle," which is to use all of the available data (Newman, 2009, p. 11). Accordingly, we echo the recommendation that the practice of dropping teams be discontinued (e.g., Cole et al., 2011; Stanley et al., 2011).

A broader implication of our field-based study is that our findings correspond to that of prior simulation research (e.g., Allen et al., 2007a; Maloney et al., 2010; Stanley et al., 2011), and, by extension, they challenge the widespread application of conventional methods for handling missing team member data. For instance, our results underscore previously noted concerns that dispersion-based variables are especially sensitive to data missingness. It seems that dispersion variables based on respondents only (SD_{WG}) may both misrepresent complete-data dispersion variables ($SD_{WG\ complete}$; Newman & Sin, 2009) and underestimate the true magnitude of dispersion variable–outcome relationships (Allen et al., 2007a, 2007b). Such a finding bolsters the dual assertion that, when within-team response rates fall below 100% (a common occurrence), the estimation of team-level properties is complicated by missing data, and the adoption of a purely ad hoc approach (i.e., listwise and pairwise deletion) to team member missingness may not be an effective solution to the underlying problem (as our findings illustrate). According to Allison (2009), for example, "there is no good excuse" for not using the more modern missing data techniques because they are now part of widely available and easily used statistical software (p. 88). This point builds upon the work of Rogelberg and Luong (1998), who recommended that "because no single perfect or correct way of assessing nonresponse bias exists, a number of procedures should be used" (p. 63). As such, we have shown how several modern missing data techniques can be used to ascertain the extent to which data missingness has plausibly affected the empirical findings of a team-level study involving data aggregation.

Of relevance to the theme conveyed in the preceding paragraph, Newman (2009) has previously suggested that authors of research involving composition-model variables use *MI* for measure-specific missing data (when *some data* on noncompliant individuals are available) and *team-level* QA_{SNP} for survey-level missing data (when *no data* on noncompliant individuals are available). Perhaps it is worth noting that having some useful data available on every nonrespondent is quite uncommon; therefore, most field-based team-level data sets would not be amenable to the *MI* approach (Modern Method 1) demonstrated in this study (a point we return to anon). Nevertheless, researchers creating composition-model variables could still use a *team-level* QA_{SNP} to correct the observed (incomplete-data) team means and standard deviations to estimate the corresponding complete-data values.

Existing Literature of Relevance

Stanley et al. (2011), among others, have recommended that researchers adopt a comprehensive study design strategy aimed at ethically maximizing within- and between-group participation rates. In this vein, scholars have identified several design factors conducive to greater survey participation (see Dillman, 2000; Rogelberg & Stanton, 2007). Nevertheless, given that a perfect response rate in field settings (i.e., real teams) is usually unattainable (Maloney et al., 2010), we sought to document how researchers have handled team-level data derived from less-than-completely represented teams. Therefore, we conducted a preliminary review of missing data reporting practices for team-based research published in the *Journal of Applied Psychology*, during the decade of 2000–2010. As our goal was simply to document the predominant missing data techniques used by researchers who analyze team-level data collected from field sites, we excluded meta-analytic studies as well as primary studies that involved either experimental methods or student samples. As a result, we identified 36 articles for review.

In conducting our review, it proved quite difficult to identify the presence and amount of missing data within teams, insofar as important details were seldom reported (cf. Maloney et al., 2010). What is apparent, however, is that most every study had missing data given that participation rates (if reported) were less than 100%. Furthermore, in making sense of the limited information reported in these studies, we deemed that the "traditional" missing data methods of listwise and pairwise deletion were the only missing data methods used in every one of these studies. This state of affairs is a concern because methodologists (e.g., Rubin, 1987), as well as the American Psychological Association (Wilkinson & the Task Force on Statistical Inference, 1999), have long discouraged the use of listwise and pairwise deletion. By the same token, through our review we found not one instance in which researchers applied modern techniques for handling missing data within teams (to include our own prior research). This is equally disconcerting, given the preponderance of evidence demonstrating the superiority of these techniques over the traditional methods of listwise and pairwise deletion (Schafer & Graham, 2002; Sinharay et al., 2001). Although our review is clearly limited in scope, it does reiterate the idea "that a substantial gap often exists between the inferential methods that are recommended in the statistical research literature and those techniques that are actually adopted by applied researchers" (Keselman et al., 1998, p. 351). To promote the use of better missing data practices, we offer several considerations that pertain to conducting multilevel (team) field research with the use of surveys.

Practical Considerations for Multilevel (Team) Research

Consider using identified surveys. For making missing data analyses more feasible to team researchers, an initial consideration is the use of identified surveys. Recently, the application of identified employee surveys has elicited a great deal of discussion (see Saari & Scherbaum, 2011, and the accompanying commentaries).

Saari and Scherbaum (2011) have argued that identified employee surveys are "uniquely beneficial for individual-level longitudinal research questions" (p. 439) yet unnecessary for team-level research. In counterpoint to this perspective, others have advocated the use of individually identified surveys as a means of providing data that can be (a) more accurately linked to one's current work unit and functional area within the organization, as well as retained for (b) exploring research questions that emerge later and (c) recombining individuals' data at higher levels of analysis to reflect changes in organizational structure (Church & Rotolo, 2011).

Another reason to use identified surveys, as implied in the present study, is for the practical purpose of performing analyses with missing data when it is not possible to compare the actual survey responses of nonrespondents with those of respondents. To effectively use either modern missing data procedure we illustrated, some type of coding system is necessary to identify from whom completed survey data have (and have not) been obtained (cf. Rogelberg et al., 2003; Rogelberg & Stanton, 2007). Therefore, it should be recognized that *some variant of identified surveys is essential* for successfully correcting the observed (incomplete) data. Because *MI* is conducted at the individual level of analysis, it is necessary to specifically identify each potential research participant as either compliant or noncompliant. Although *team-level* QA_{SNP} entails adjusting the observed team-level means and standard deviations, the equations used as part of the correction process require the precise percentage of nonparticipation within each team.

Consider including auxiliary variables. To facilitate *MI*, another strategy is to include auxiliary variables (Graham, 2009) linked to individual team members. An auxiliary variable represents a construct outside of the focal theoretical model, yet one that is used for the imputation process. Good auxiliary variables (see Collins, Schafer, & Kam, 2001) can be instrumental for obtaining more accurate imputations, thereby reducing bias and improving the precision of parameter estimates (Allison, 2009). Practically speaking, the selection of auxiliary variables should stem from the anticipated strength of their associations with the focal research variables for which some data are missing (Little, 1995). Toward this end, researchers could incorporate different sources of data (e.g., survey data, demographic data obtained from individuals' personnel records, etc.) when identifying potential auxiliary variables. For example, we were able to use an array of pertinent auxiliary variables with data on individuals obtained from external sources. Alternatively, if individual-level data are not available, basic information about every team (e.g., team size, team member tenure, types of tasks undertaken, etc.) could be gathered and used in the imputation process (Maloney et al., 2010). Readers interested in the application of auxiliary variables may wish to consult Collins et al. (2001) and Graham (2009).

Provide transparent information. To paraphrase Maloney and colleagues (2010, pp. 295–296), more transparent and detailed reporting of within-group participation rates will promote a better understanding of the extent to which members' nonresponse impacts the empirical findings of a study. Likewise, we join other researchers in calling for, at a minimum, the reporting of average within- and between-team response rates, as well as that of how within-team response rates are distributed (Allen et al., 2007b; Newman & Sin, 2009). This should include information about teams' size (e.g., mean, min, and max), range of within-team participation, and some distributional data about how team member nonresponse (i.e., missing data) varied among the sampled teams.

Use all available data and apply modern missing data techniques. It is likely that some data sets will have teams with high representation and others with low representation. Rather than eliminating teams with low levels of representation (prior to analyzing the data), a better option is to apply modern missing data techniques and sensitivity tests to evaluate the impact of systematic nonresponse (d_{miss}) on the aggregated team-level variables. Moreover, when applying SNPs to composition-model variables, it is reasonable for researchers to incorporate nonlocal, yet plausible, estimates of d_{miss}. The value of doing so stems from the reality that there is no single way of determining the precise extent to which the actual d_{miss}, pertaining to the specific variables for which data are missing, is limiting the generalizability of findings (Rogelberg & Luong, 1998). Thus, by using varying (moderate to extreme) values of d_{miss}, researchers can easily ascertain how sensitive their results are to missing data. If the sensitivity analyses yield a pattern of similar results, researchers can be more confident that the mixing of high- and low-representation teams does not undermine the accuracy of hypothesis testing. If the sensitivity results reveal inconsistencies, then the accuracy of the estimated relationships (i.e., internal validity) could be called into question. In such a situation, authors should stipulate what readers should *not* infer from the reported results and thereby offer only those generalizations that are reasonable (Rogelberg & Luong, 1998).

Conclusion

Aggregated survey data are prevalent in field research on work teams. As such, it is important to understand how missing data complicate the estimation of relationships between aggregated (i.e., composition-model variables) team-level constructs and outcomes (e.g., Allen et al., 2007a; Timmerman, 2005). Our exploration provides initial evidence that team-level phenomena may be systematically reflected in team members' voluntary survey (non)completion, and it offers original insight into the nature and magnitude of such links. Future team research will benefit by accounting for any plausible systematic effects associated with missing data. Toward that end, we have illustrated a comprehensive approach for dealing analytically with systematic data missingness in teams. It is hoped that the methods shown will serve as exemplars for team researchers wishing to examine bottom-up constructs based on consensus- and/or dispersion-composition models.

References

Allen, N. J., Stanley, D. J., Williams, H. M., & Ross, S. J. (2007a). Assessing dissimilarity relations under missing data conditions: Evidence from computer simulations. *Journal of Applied Psychology, 92,* 1414–1426. doi:10.1037/0021-9010.92.5.1414

Allen, N. J., Stanley, D. J., Williams, H. M., & Ross, S. J. (2007b). Assessing the impact of nonresponse on work group diversity effects. *Organizational Research Methods, 10,* 262–286. doi:10.1177/1094428106/294731

Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: Sage.

Allison, P. D. (2009). Missing data. In R. E. Millsap & A. Maydeu-Olivares (Eds.), *The Sage handbook of quantitative methods in psychol-*

ogy (pp. 72–89). Thousand Oaks, CA: Sage. doi:10.4135/9780857020994.n4

Becker, T. E. (2005). Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations. *Organizational Research Methods, 8,* 274–289. doi:10.1177/1094428105278021

Biemann, T., Cole, M. S., & Voelpel, S. (2012). Within-group agreement: On the use (and misuse) of r_{WG} and $r_{WG(j)}$ in leadership research and some best practice guidelines. *The Leadership Quarterly, 23,* 66–80. doi:10.1016/j.leaqua.2011.11.006

Black, J. G., Hyland, P. K., & Rutigliano, P. (2011). Realizing the promise and minimizing the perils of identified surveys: Reports from the field. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 4,* 462–467. doi:10.1111/j.1754-9434.2011.01374.x

Bliese, P. D., & Halverson, R. R. (1998). Group consensus and psychological well-being: A large field study. *Journal of Applied Social Psychology, 28,* 563–580. doi:10.1111/j.1559-1816.1998.tb01720.x

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,* 234–246. doi:10.1037/0021-9010.83.2.234

Chan, D. (2011). Advances in analytical strategies. In S. Zedeck (Ed.), APA handbook of industrial and organizational psychology: Vol. *1*. *Building and developing the organization* (pp. 85–113). Washington, DC: American Psychological Association.

Church, A. H., & Rotolo, C. T. (2011). Revisiting the great survey debate: Aren't we past that yet? *Industrial and Organizational Psychology: Perspectives on Science and Practice, 4,* 455–459. doi:10.1111/j.1754-9434.2011.01372.x

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cole, M. S., Bedeian, A. G., Hirschfeld, R. R., & Vogel, B. (2011). Dispersion-composition models in multilevel research: A data-analytic framework. *Organizational Research Methods, 14,* 718–734. doi:10.1177/1094428110389078

Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6,* 330–351. doi:10.1037/1082-989X.6.4.330

De Cremer, D., & van Knippenberg, D. (2002). How do leaders promote cooperation? The effects of charisma and procedural fairness. *Journal of Applied Psychology, 87,* 858–866. doi:10.1037/0021-9010.87.5.858

De Jong, B. A., & Dirks, K. T. (2012). Beyond shared perceptions of trust and monitoring in teams: Implications of asymmetry and dissensus. *Journal of Applied Psychology, 97,* 391–406. doi:10.1037/a0026483

Dillman, D. A. (2000). *Mail and Internet surveys: The tailored design method*. New York, NY: Wiley.

Dovidio, J. F., Piliavin, J. A., Schroeder, D. A., & Penner, L. A. (2006). *The social psychology of prosocial behavior*. Mahwah, NJ: Erlbaum.

Flynn, F. J. (2011). Give and take: Psychological mindsets in conflict. In D. De Cremer, R. van Dick, & J. K. Murnighan (Eds.), *Social psychology and organizations* (pp. 231–250). New York, NY: Routledge.

Froelich, J. M. (2011). Identifying the ethical (unethical) undercurrent of identified surveys. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 4,* 476–478. doi:10.1111/j.1754-9434.2011.01377.x

Goldstein, N. J., & Cialdini, R. B. (2011). Managing normative influences in organizations. In D. De Cremer, R. van Dick, & J. K. Murnighan (Eds.), *Social psychology and organizations* (pp. 67–86). New York, NY: Routledge.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60,* 549–576. doi:10.1146/annurev.psych.58.110405.085530

Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: Description and an illustration. *Public Opinion Quarterly, 64,* 299–308. doi:10.1086/317990

Hackman, J. R. (1992). Group influences on individuals in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (Vol. *3,* 2nd ed., pp. 199–267). Palo Alto, CA: Consulting Psychologists Press.

Hirschfeld, R. R., & Bernerth, J. B. (2008). Mental efficacy and physical efficacy at the team level: Inputs and outcomes among newly formed action teams. *Journal of Applied Psychology, 93,* 1429–1437. doi:10.1037/a0012574

Hollenbeck, J. R., Beersma, B., & Schouten, M. E. (2012). Beyond team types and taxonomies: A dimensional scaling conceptualization for team description. *Academy of Management Review, 37,* 82–106.

Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., . . . Levin, J. R. (1998). Statistical practices of educational researchers: An analysis of their ANOVA, MANOVA, and ANCOVA analyses. *Review of Educational Research, 68,* 350–386.

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 3–90). San Francisco, CA: Jossey-Bass.

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11,* 815–852. doi:10.1177/1094428106296642

Levi, D. (2007). *Group dynamics for teams* (2nd ed.). Thousand Oaks, CA: Sage.

Levy, D. A., & Nail, P. R. (1993). Contagion: A theoretical and empirical review and reconceptualization. *Genetic, Social, and General Psychology Monographs, 119,* 233–284.

Lindell, M. K., & Brandt, C. J. (2000). Climate quality and climate consensus as mediators of the relationship between organizational antecedents and outcomes. *Journal of Applied Psychology, 85,* 331–348. doi:10.1037/0021-9010.85.3.331

Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of American Statistical Association, 90,* 1112–1121. doi:10.1080/01621459.1995.10476615

London, M., & London, M. (2007). *First-time leaders of small groups: How to create high-performing committees, task forces, clubs, and boards*. San Francisco, CA: Jossey-Bass.

Maloney, M. M., Johnson, S. G., & Zellmer-Bruhn, M. E. (2010). Assessing group-level constructs under missing data conditions: A Monte Carlo simulation. *Small Group Research, 41,* 281–307. doi:10.1177/1046496410363829

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review, 26,* 356–376.

Mathieu, J., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future. *Journal of Management, 34,* 410–476. doi:10.1177/0149206308316061

Meade, A. W., & Eby, L. T. (2007). Using indices of group agreement in multilevel construct validation. *Organizational Research Methods, 10,* 75–96. doi:10.1177/1094428106289390

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1978). *Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. Retrieved from http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html

Nesterkin, D. A., & Ganster, D. G. (2012). The effects of nonresponse rates on group-level correlations. *Journal of Management*. Advance online publication. doi:10.1177/0149206311433853

Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods, 6,* 328–362. doi:10.1177/1094428103254673

Newman, D. A. (2009). Missing data techniques and low response rates: The role of systematic nonresponse parameters. In C. E. Lance & R. J. Vandenberg (Eds.), *Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences* (pp. 7–36). New York, NY: Routledge.

Newman, D. A., & Sin, H.-P. (2009). How do missing data bias estimates of within-group agreement? Sensitivity of SD_{WG}, CV_{WG}, $r_{WG(J)}$, $r_{WG(J)}^{*}$, and ICC to systematic nonresponse. *Organizational Research Methods, 12,* 113–147. doi:10.1177/1094428106298969

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research, 74,* 525–556. doi:10.3102/00346543074004525

Preacher, K. J., Rucker, D. D., MacCallum, R. C., & Nicewander, W. A. (2005). Use of the extreme groups approach: A critical reexamination and new recommendations. *Psychological Methods, 10,* 178–192. doi:10.1037/1082-989X.10.2.178

Reich, T. C., & Hershcovis, M. S. (2011). Interpersonal relationships at work. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology*: Vol. *3. Maintaining, expanding, and contracting the organization* (pp. 223–248). Washington, DC: American Psychological Association.

Rizzuto, T. E., Mohammed, S., & Vance, R. J. (2011). Marching in-step: Facilitating technological transitions through climate consensus. *Computers in Human Behavior, 27,* 185–194. doi:10.1016/j.chb.2010.07.035

Roberson, Q. M., Sturman, M. C., & Simons, T. L. (2007). Does the measure of dispersion matter in multilevel research? A comparison of the relative performance of dispersion indexes. *Organizational Research Methods, 10,* 564–588. doi:10.1177/1094428106294746

Rogelberg, S. G., Conway, J. M., Sederburg, M. E., Spitzmüller, C., Aziz, S., & Knight, W. E. (2003). Profiling active and passive nonrespondents to an organizational survey. *Journal of Applied Psychology, 88,* 1104–1114. doi:10.1037/0021-9010.88.6.1104

Rogelberg, S. G., & Luong, A. (1998). Nonresponse to mailed surveys: A review and guide. *Current Directions in Psychological Science, 7,* 60–65. doi:10.1111/1467-8721.ep13175675

Rogelberg, S. G., Luong, A., Sederberg, M. E., & Cristol, D. S. (2000). Employee attitude surveys: Examining the attitudes of noncompliant employees. *Journal of Applied Psychology, 85,* 284–293. doi:10.1037/0021-9010.85.2.284

Rogelberg, S. G., & Stanton, J. M. (2007). Introduction: Understanding and dealing with organizational survey nonresponse. *Organizational Research Methods, 10,* 195–209. doi:10.1177/1094428106294693

Rosh, L., Offermann, L. R., & Van Diest, R. (2012). Too close for comfort? Distinguishing between team intimacy and team cohesion. *Human Resource Management Review, 22,* 116–127. doi:10.1016/j.hrmr.2011.11.004

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley & Sons. doi:10.1002/9780470316696

Saari, L. M., & Scherbaum, C. A. (2011). Identified employee surveys: Potential promise, perils, and professional practice guidelines. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 4,* 435–448. doi:10.1111/j.1754-9434.2011.01369.x

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7,* 147–177. doi:10.1037/1082-989X.7.2.147

Sinharay, S., Stern, H. S., & Russell, D. (2001). The use of multiple imputation for the analysis of missing data. *Psychological Methods, 6,* 317–329. doi:10.1037/1082-989X.6.4.317

Spitzmüller, C., Glenn, D. M., Barr, C. D., Rogelberg, S. G., & Daniel, P. (2006). "If you treat me right, I reciprocate": Examining the role of exchange in organizational survey response. *Journal of Organizational Behavior, 27,* 19–35. doi:10.1002/job.363

Spitzmüller, C., Glenn, D. M., Sutton, M. M., Barr, C. D., & Rogelberg, S. G. (2007). Survey nonrespondents as bad soldiers: Examining the relationship between organizational citizenship and survey response behavior. *International Journal of Selection and Assessment, 15,* 449–459. doi:10.1111/j.1468-2389.2007.00403.x

Stajkovic, A. D., Lee, D., & Nyberg, A. J. (2009). Collective efficacy, group potency, and group performance: Meta-analyses of their relationships, and test of a mediational model. *Journal of Applied Psychology, 94,* 814–828. doi:10.1037/a0015659

Stanley, D. J., Allen, N. J., Williams, H. M., & Ross, S. J. (2011). Examining workgroup diversity effects: Does playing by the (group-retention) rules help or hinder? *Behavior Research Methods, 43,* 508–521. doi:10.3758/s13428-010-0053-9

Thatcher, S. M. B., & Patel, P. C. (2011). Demographic faultlines: A meta-analysis of the literature. *Journal of Applied Psychology, 96,* 1119–1139. doi:10.1037/a0024167

Timmerman, T. A. (2005). Missing persons in the study of groups. *Journal of Organizational Behavior, 26,* 21–36. doi:10.1002/job.306

van Bruggen, G. H., Lilien, G. L., & Kacker, M. (2002). Informants in organizational marketing research: Why use multiple informants and how to aggregate responses. *Journal of Marketing Research, 39,* 469–478. doi:10.1509/jmkr.39.4.469.19117

Wilkinson, L., & the Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54,* 594–604. doi:10.1037/0003-066X.54.8.594

Zajonc, R. B. (1965). Social facilitation. *Science, 149,* 269–274. doi:10.1126/science.149.3681.269

Received July 12, 2011
Revision received December 28, 2012
Accepted January 14, 2013 ■