That Laboratory-Derived Findings Generalize to Work Teams:
A Search for the Supporting Evidence

Adam C. Morgan
University of Technology, Sydney (UTS)
adam.morgan@uts.edu.au

Paper presented at the 2013 Academy of Management Annual Meeting, Orlando, FL
That Laboratory-Derived Findings Generalize to Work Teams: 
A Search for the Supporting Evidence

ABSTRACT

At the start of their landmark review of the work group/team literature, Gist, Locke and Taylor (1987) issued the following warning: “the reader is cautioned that although we assumed generalizability from lab to field settings in our presentation, it is an assumption and we lack the evidence to support this as a conclusion” (p. 239). Leveraging off this statement by Gist et al. (1987), this study investigates the extent to which any such evidence has emerged since 1987, and does so by an examination of all team-related articles published between 1987 and 2011 in five management/OB journals; Academy of Management Journal, Administrative Science Quarterly, Journal of Applied Psychology, Journal of Management, and Organizational Behaviour and Human Performance/Decision Processes. The results of this study are that the point made by Gist et al. (1987) about the lack of evidence is one that recurs frequently in the teamwork literature, yet despite this continuing acknowledgement of a significant gap, little research attention has been paid to exploring the evidential basis for generalizability with reference to teams. This paper presents and discusses the few articles that were found to relate to the issue. Some broad observations arising from this study are also presented and discussed, and some future research directions are outlined.

Keywords: 
Work Teams; Lab to Field Generalizability; Between-Setting Research
That Laboratory-Derived Findings Generalize to Work Teams:
A Search for the Supporting Evidence

For more than two decades, the Journal of Management has published periodic reviews of the teamwork literature. Gist, Locke, and Taylor (1987) conducted the first such review for the journal, with Mathieu, Maynard, Rapp, and Gilson (2008) providing the most recent to date. In between, reviews have been offered by Bettenhausen (1991) and Cohen and Bailey (1997). Each of these reviews covers empirical studies carried out in the field. Each review also covers laboratory-based studies, with the reason for their inclusion given at the start of the paper.

In the first of this review series, Gist et al. (1987) provided the following comment regarding their decision to include laboratory studies:

The studies reviewed were conducted in both laboratory and field settings. This review draws no firm conclusions on the generalizability of results from lab to field. Such determinations require extensive analyses of all studies on particular topics, whereas this review is limited to a particular chronological period. However, a recent collection of literature reviews on several organizational behavior and human resource management topics (Locke, 1986) suggests that similar findings result from research conducted in both lab and field settings. These results led to our decision to assume generalizability from lab to field in the articles reviewed. However, this assumption may not typify small group research because the effects of a larger organizational environment may interact with group characteristics to yield findings that differ from those produced by a laboratory work group in isolation. Thus, the reader is cautioned that although we assumed generalizability from lab to field
settings in our presentation, it is an assumption and we lack the evidence to support this as a conclusion (pp. 238-9).

In the next instalment of the series, Bettenhausen (1991) also chose to review laboratory-based research, which was explained as follows:

Practical issues are often best addressed with an experimental design; moreover, depending on the issue, findings using student groups in manipulated settings often generalize to organizational settings as well or better than findings from intact groups that are frequently confounded by unique, unmeasured contextual factors. Insight into the functioning of work groups comes from the juxtaposition of the findings and conclusions of research conducted in a variety of settings. To capture these insights, this review includes conceptual as well as empirical work and does not censor work merely because it was conducted in unique contexts, such as sports teams or behavioral labs, or from unconventional disciplines, such as psycholinguistics or architecture (p. 346).

In both Gist et al. (1987) and Bettenhausen (1991) generalizability is assumed, notwithstanding the caveats expressed in both studies. However, when Cohen and Bailey (1997) conducted the next review in this series, the use of laboratory-based research was reduced from their review of the literature, though not excluded entirely. Their rationale for doing so was given as follows:

Our focus for this review is studies of teams in organizational settings…. We chose this focus for three major reasons. First, the findings from teams performing
real tasks in organizational settings can more readily be generalized to the world of work. The findings from studies of undergraduate psychology or business students are much less likely to apply to practicing research and development managers, blue collar workers, or executives. Second, organizational features external to the team can be extremely important determinants of effectiveness, yet they are rarely examined in laboratory settings (Hackman, 1987).... Third, we wanted to review a coherent body of work. Thus, we are willing to sacrifice the rigor of the experimental laboratory to deal with the confounds of the real world in the hope that the findings we identify can be used to guide management practice and to improve effectiveness.

However, we do review conceptual and theoretical articles to help us understand the factors that contribute to effectiveness of teams. We also occasionally present findings from experiments done in the laboratory. We make these exceptions for newer areas of research such as group cognition in which more conceptual than empirical work has been done and few field studies exist (p. 240).

In the most recent instalment of the series, where the years 1997-2007 are covered, Mathieu et al. (2008) decided to include laboratory-based research in their review, effectively reinstating the position of assumed generalizability. Unlike previous reviews in the series, however, a detailed introductory statement on their decision was not given. The authors simply noted the following:

Whereas we include research that has been conducted in laboratory or simulated work environments, the target of generalizations for us are teams in organizations
rather than social groups, sports teams, or collectives that operate in other contexts

[italics in original] (p. 411).

From the quotations presented above, it is clear that laboratory to field generalizability is an “issue” within the work group/team literature, with opinion on the issue varying somewhat. For Gist et al. (1987), Bettenhausen (1991), and Mathieu et al. (2008), generalizability was not deemed to be an issue leading to the exclusion of laboratory-based research. Whereas Gist et al. (1987), Bettenhausen (1991), and Mathieu et al. (2008) assume generalizability, albeit with varying degrees of confidence, the position held by Cohen and Bailey (1997) appears to differ from the others. Cohen and Bailey (1997), argue that bona fide work team samples are more readily generalizable, although they do not go so far as to discount lab results altogether.

Importantly, it seems that none of these authors cites any empirical evidence to support their stance on the issue of generalizability (e.g., empirical research where the results deriving from laboratory groups and work teams have been statistically compared). Whatever their degree of reservation or endorsement regarding the issue of generalizability, it is striking that the question of supporting evidence for their respective positions remains largely unexplored. The exception is Gist et al. (1987), who in their introductory remarks, assume generalizability from lab to field settings but at the same time explicitly acknowledge that it is just that, an assumption, and that they “lack the evidence to support this as a conclusion” (p. 239). It is now 25 years since Gist et al. (1987) made this crucial reference to the keyword “evidence.” What has happened since, and what is the situation now?

Judging by the quotations presented above, Gist et al.’s (1987) evidence may still be lacking; it is highly likely that had supporting evidence been available at the time of any or all of
the studies under discussion here, this evidence would have been cited by the authors. Drawing this conclusion from a sample of just four papers, is not, however, sufficiently convincing. What is needed is a more thorough examination of the teamwork literature to test whether or not the evidence does exist in the form of published research.

The search for the evidence involves a comprehensive review of the teamwork literature. Using the publication date of Gist et al.’s (1987) initial study as the starting point, this review will attempt to identify and examine all studies involving comparisons made between laboratory setting groups and bona fide work teams published within a key set of management/organizational science journals over the last 25 years. It will also examine the comments made by authors using laboratory setting samples to see whether any evidence is cited. Close examination of the teamwork literature shows that many team researchers have used groups operating in laboratory settings in their investigations, with the intent being to generalize their findings to bona fide work teams. It also shows that various articles are cited to support their choice of sample. Of interest to this study is what kind of evidence is provided.

**Points of Clarification**

Prior to outlining the methods employed in this study and the results found, it is important to make five clarifying points. First, this paper uses the word “team” as the primary term, but draws no distinction between this term and “group” in an organizational context (e.g., work groups, groups in organizations, organizational groups). Consistent with most other authors (e.g., Cohen & Bailey, 1997; Guzzo & Dickson, 1996; Guzzo & Shea, 1992; Ilgen, 1999; Tannenbaum, Beard, & Salas, 1992), these terms of team and group (in an organizational context) are treated as interchangeable.
Second, a team is considered to be a social entity, consistent with the following definition offered by Cohen and Bailey (1997): “A team is a collection of individuals who are interdependent in their tasks, who share responsibility for outcomes, who see themselves and who are seen by others as an intact social entity embedded in one or more larger social systems (for example, business unit or the corporation), and who manage their relationships across organizational boundaries” (p. 241).

Third, it is recognized that the broad issue of laboratory to field generalizability has been addressed by a number of authors over the years (e.g., Anderson, Lindsay, & Bushman, 1999; Berkowitz & Donnerstein, 1982; Cardy, 1991; Dipboye & Flanagan, 1979; Dobbins, Lane, & Steiner, 1988; Driskell & Salas, 1992; Gordon, Slade, & Schmitt, 1986, 1987; Greenberg, 1987; Ilgen, 1986; Locke, 1986; Marks, 2000; Mook, 1983). Further, some of this literature makes reference to teams and teamwork (e.g., Driskell & Salas, 1992; Ilgen, 1986; Marks, 2000). This literature, however, is conceptual, and does not specifically deal with the issue of evidence with regards to team-related lab to field generalizability. This paper seeks to focus on this identified issue.

Fourth, for the purpose of this study, students working together on assignments in classroom settings are excluded. Although these classroom groups are operating within the field, in so far as they are naturally occurring groups and not formed for research purposes, they constitute a different type of “field” to that of the workplace. Further, groups operating in other types of nonworkplace field settings, such as sport teams, are excluded from this study; like student groups in the classroom context, they are treated as an “other” category.

Fifth and finally, it is acknowledged that the evidence being sought could take a variety of forms. It might include a comparative study, such as those presented in Locke’s (1986) edited
book: *Generalizing from Laboratory to Field Settings*. It might take the form of a meta-analysis. It might equally take the form of research where work and lab setting samples have been used and the derived results compared; either in the same paper or in additional ones involving replications.

**METHODS**

As this study sought to examine team-related articles over a 25-year period, the number of journals included in this review was restricted to a relatively small, yet reputable, set. This set comprised the following five journals: *Academy of Management Journal* (AMJ), *Administrative Science Quarterly* (ASQ), *Journal of Applied Psychology* (JAP), *Journal of Management* (JOM), and *Organizational Behaviour and Human Decision Processes* (OBDHP, for the years 1985-2011 and *Organizational Behaviour and Human Performance*, for the years 1987-1994).

Following the identification of the journals, team-related papers for the intended review were identified. This was done by first using the PsycINFO database to identify all articles that, within the five targeted journals for the years 1987-2011, had one or more of the following “wildcard” terms in either their title or abstract: team, group, quality circle, or crew. A manual culling of the 1384 identified “hits” followed, eliminating those that were reviews, theoretical pieces, editorials, book reviews, errata, and those having little or no direct relevance to the study of work teams. This final culling criterion was required as, by including the search term of “group,” many articles were identified that were clearly not team-related, such as those where group was used in a different context (e.g., relating to control groups, groups of organizations, specific groups of workers). This final culling criterion was also required as the degree of “team-relatedness” varied considerably in the identified articles. At one end of the continuum,
there were articles that were clearly related to the study of work teams. Apart from having “team” or “work group” in their title, these articles also usually contained a positioning statement in their introductory paragraphs related to teamwork (e.g., “organizations increasingly rely on teams…”). Towards the other end of the continuum, there were articles that had their primary focus on topics other than work teams (e.g., leadership, relational demography, social networks), with little or no positioning of teamwork in their introductions. To cull these articles having teams as a secondary focus, a holistic judgement was used, taking into account the authors’ positioning statements, the presence or absence of “team” or “work group” in the title, and when available, the author-supplied keywords. Three hundred and forty-eight papers remained after this culling exercise.

A search through these 348 articles was then undertaken looking for studies where, within the one paper, data were collected from both lab and work settings. Those found were then examined in detail. Any articles that sought to replicate research from a lab setting to a work setting, or vice versa, were also examined, as were any meta-analytic reviews that investigated between-setting differences. Finally, all papers using lab setting samples were examined with reference to their comments surrounding the issue of generalizability. Particular attention was paid to the sources cited in defense of sample choice made in these papers, and whether these sources might help to illuminate the generalizability assumption identified by Gist et al. (1987); in other words, whether these papers cited sources in their defense that were based on comparative research.
RESULTS

Multiple Setting Studies

Of the 348 articles examined in detail as part of this study, the total number that collected data from both lab and work settings was only two: Chen, Sharma, Edinger, Shapiro, and Farh (2011) and Earley and Mosakowski (2000). In terms of providing evidence for or against lab to field generalizability, it is difficult to draw strong conclusions from their findings, as neither of these studies sought to directly investigate this issue by way of statistically comparing sample-specific results. It is made even more difficult as within each study there are methodological differences that limit the direct comparison of the samples’ results. In Earley and Mosakowski’s (2000) research, for example, their lab setting study (study 2) involved quantitative analyses, whereas their work setting study (study 1) was a qualitative study based on data collected from five work teams; it involved, according to Earley and Mosakowski (2000), “direct observations of team meetings, company records of demographic information, and open-ended and structured onsite interviews with key team personnel” (p. 30).

In Chen et al.’s (2011) research, one of the dependent variables, teamwork behaviour, was measured differently in their work and lab samples. For their work sample, team members’ actual behaviour was rated by the team leader. In contrast, the teamwork behavior measure in their lab sample was based on the assessment of how the student would behave; the students completed items such as “I would probably work to make sure the Task Force succeeds” (Chen et al., 2011, p. 546). Many other differences existed in how data in the two settings were collected.

Despite the methodological differences in their samples, comments are made in both studies regarding the comparison of their lab and field findings. Earley and Mosakowski (2000)
suggest that their results were sufficiently similar to provide demonstrated generalizability. “The field observations from study 1 converged with the findings from [study 2], demonstrating the mundane realism and generalizability of these results” suggest Earley and Mosakowski (2000, p. 47). Chen et al. (2011) found that a number of their hypotheses were supported in both of their samples; in other words, their results in each sample were sufficiently strong to be significant. However, they also found that some of the hypotheses supported in their lab study, were unable to be supported in their field study (e.g., those concerning behaviour). This led to the authors to reflect on this apparent discrepancy and to offer the following comment in the section entitled Limitations and Future Research Directions:

The findings in our field study (Study 2) were not as strong as those of our experimental study (Study 1). One explanation for these differences in findings may be that the field study, relative to the lab study, may have lacked sufficient statistical power to detect our hypothesized effects. Another possible reason is that we measured behavioral intentions in the lab study but ratings of actual behaviors in the field sample. Hence, it could be that the direct and mediated influences of team stimuli were stronger when considering behavioral intentions, which are more proximal to team stimuli and motivational states than actual behaviors (cf. Chen & Kanfer, 2006). Yet other possibilities for the inconsistency in findings might be the diversity in team types represented in the field sample (i.e., our hypothesized effects might be stronger in some types of teams than others), as well as the low ICC(2) values we obtained for measures of team stimuli in Study 2, which could have attenuated our findings (cf. Bliese, 2000). Thus, although we attempted to maximize internal validity in Study 1 and external validity in Study
2, clearly, replications that take these measurement and sampling differences into account are needed (p. 555).

These comments by Chen et al. (2011) represent a significant insight into the question of generalizability, suggesting as they do, a number of possible reasons for why the sample differences existed.

**Replications**

Along with the search for multi-setting articles, this study also searched for papers where an effort was made to replicate findings from a lab setting to a work setting, or vice versa. Of the 348 articles examined, only one could be found: Gersick (1989). Gersick’s paper reports the results of a laboratory study that, among other aims, attempted to compare findings attained from research published the year before (Gersick, 1988). More specifically, it attempts to examine if Gersick’s (1988) observations of a punctuated equilibrium with a midpoint transition discovered in field groups could also be found in laboratory groups. This laboratory research by Gersick (1989) involved the recording and analysis of the verbal communication taking place between members of groups working together on a novel task under time constraints. Eight groups were observed in this research, each comprised of MBA students.

Following the analysis of the collected data, Gersick (1989) included a section entitled “A Field-Laboratory Comparison” within the paper’s discussion. In this section, Gersick (1989) makes a number of comments regarding the differences and similarities of the two studies’ results (i.e., the laboratory study and the field research published in the previous year). According to Gersick (1989):
There were some differences between the dynamics observed in this laboratory study and the dynamics in the naturally occurring groups. In the field, groups met on several occasions. They prepared homework and communicated with independently acting stakeholders between meetings, and they often extended their midpoint meetings beyond their usual time limits in order to complete a key piece of work. Because of the laboratory design, the groups described in this study lacked the flexibility to extend their time spans, so it was not possible to see comparable changes in groups’ routines or external interactions. It was not possible to identify a clearly bounded “transition meeting” with a set of transition accomplishments. Because of these restrictions, too, laboratory transitions were more temporally compact and less complex than some of the transitions observed in the field. More important, the laboratory setting did not simulate team-supervisor relations very well, thus restricting the generalizability of results dealing with that facet of team development (p. 305).

Gersick (1989) also noted that:

A final difference has to do with the shortness of the time span for the laboratory teams’ work. The fact that the lab groups made many more time comments overall than the field groups and the increased frequency of all the lab groups’ time comments toward the end of their meeting hour suggest that the sharper a group’s time constraints, the more frequently participants will pay explicit attention to time and pacing…. Since the midpoint effect was first discovered in the field and then observed in the lab, that particular finding appears to be robust.
However, it is important to be cautious about generalizing pacing patterns back to the field wholesale (p. 305).

Finally, Gersick (1989) suggests that:

Notwithstanding these differences, the central features of the midpoint transition did emerge strongly in the laboratory study: groups paid special attention to time at the midpoint transition of their time spans.… If laboratory task groups can display a punctuated equilibrium pattern with a midpoint transition, we can use laboratories to see how various interventions cause groups to depart from that baseline. Such settings appear to be promising sites for further research (p. 305).

These comments made by Gersick (1989) are interesting because they appear to suggest that generalizability was not observed in all areas, but rather only in particular parts. At the same time, it should be noted that Gersick’s (1988) field research was based on a small number of groups (eight in total), and relied mainly on the qualitative analysis of verbal communication. Further, Gersick’s (1988) field study sample was not entirely comprised of work teams; it contained five work teams and three groups of MBA students working on graded assignments. Nevertheless, it is an important—and a very rare—piece of research that sought to make some much needed between-setting comparisons.

**Meta-Analyses**

The next place where this study searched for evidence was in the published meta-analytic studies. Although no study was found like that conducted by Anderson et al. (1999; where lab
versus field was investigated in a range of social psychological studies), six instances were found where the lab-field issue was investigated as a moderating variable in the meta-analysis. These were the studies conducted by Bell (2007), Bell Villado, Lukasik, Belau, and Briggs (2011), DeChurch and Mesmer-Magnus (2010), Horwitz and Horwitz (2007), Kleingeld, van Mierlo and Arends (2011), and Thatcher and Patel (2011). A summary of the relationships examined for each study, and the results associated with the lab versus field setting moderation effect can be found in Table 1. Overall, these studies show “mixed” results. On one hand there are studies that found no between-setting moderation effect (e.g., Horwitz & Horwitz, 2007 and Kleingeld et al., 2011). On the other, however, most did discover various between-setting differences. Bell (2007), for example, found that the study setting was a strong moderator when examining various team composition → team performance relationships (e.g., using compositional variables, such as conscientiousness and extraversion). Further, Bell (2007) reported that many of these relationships were stronger in the field than in the lab setting. This is the same as what Bell et al. (2011) and DeChurch and Mesmer-Magnus (2010) discovered; both sets of authors reported results where the field setting relationships were stronger than those conducted in the lab. Thatcher and Patel (2011), though, found the opposite. They reported that in their research “the strength of the effects for the lab studies was consistently stronger than that of field studies [they examined] (p. 1127). Nevertheless, Thatcher and Patel’s (2011) results again point to differences existing between lab and field studies, which is the substantive point. The dates of these six meta-analyses should also be noted, for they have all been published quite recently. Although still few in number at this point, these studies taken together do appear to indicate that the lab versus field issue remains current and requires further investigation.
Sources Cited

The final place where this study searched for Gist et al.’s (1987) evidence was in the sources cited by those authors using lab samples. Of the 198 articles found to be using lab samples, most of them cited sources in relation to their sample choice and the issue of generalizability. Commonly cited sources, however, were not discovered. Berkowitz and Donnerstein’s (1982) and Ilgen’s (1986) pieces on generalizability were the closest to commonly cited sources, with twelve and ten citations, respectively. Driskell and Salas (1992) and Locke (1986) were the next most cited, with nine citations each. This is an interesting “top 4,” as the first three are nonempirical pieces, and the last is an edited book reviewing laboratory to field generalizability research on a range of organizational topics other than teams.

Inspection of the less commonly cited sources also failed to uncover papers providing strong evidence for generalizability. In most instances, the sources cited related to studies where evidence of generalizability was found in areas away from work teams, but with inferences drawn back to them. Usually, an assumption was being made by the author(s) that, as laboratory to field generalizability has been found in other applied domains, the same should apply to the study of teams. Some of the sources cited in defense of generalizability include Anderson et al. (1999), Bandura (1997), Bono and Ilie (2006), and Locke and Latham (1990).

There were also instances where the source cited related to prior research that was also conducted with a lab sample. In these instances, the authors appear to be suggesting that their sample choice is permissible due to the fact that other researchers have used similar samples in the past, usually coupled with similar experimental tasks. But the comparison is being made to the method, rather than the results.
DISCUSSION

It is an assumption and we lack the evidence to support this as a conclusion (Gist et al., 1987, p. 239).

These were the qualifying remarks made by Gist et al. (1987) at the start of their team-related review, and the genesis of this study. The conclusion they are referring to is that, with reference to the topic of organizational teams, the results deriving from laboratory research are generalizable to the field. That is, what is found to occur in the laboratory is externally valid, and therefore applicable to how teams and their members operate together in organizational settings. This conclusion, though, was acknowledged to be an assumption by Gist et al. (1987), for they lacked the evidence to support it. All they had to leverage off was the support for laboratory to field generalizability in a number of other organizational domains; but not that concerning organizational teams. This “lack of evidence” statement was made more than 25 years ago. What has been done on the issue of generalizability since then? To what extent does their “evidence” now exist? The purpose of this study was to discover answers to these questions, by conducting a review of the teamwork literature; a somewhat extensive review covering all team-related articles in five management/OB journals over a 25-year period.

Based on the review conducted for this study, Gist et al.’s (1987) evidence still appears to be lacking. This conclusion is largely based on the lack of research activity discovered. Of the hundreds of articles examined in this review, the number contributing directly to the issue—in terms of offering between-setting/sample comparisons—was less than ten. This is a surprising finding, as it does not reflect the number of times the issue is raised by team researchers in the
literature, which easily totals several hundred\textsuperscript{1}. It is an issue that is widely acknowledged but, at least to date, little explored.

This leaves the question of whether it is an issue worthy of serious attention by researchers. Judging by the large number of generalizability-related comments raised by authors in the teamwork literature reviewed, coupled by the fact that many stress the need for their research to be replicated in the field to specifically address the issue, the answer must be yes.

The results of the review conducted for this study also raise the question of why there has been little empirical attention paid to the issue to date. One explanation may lie with the motivation of team researchers to avoid entering into the lab versus field debate. Some may have even been influenced by John Campbell, who, in his chapter included in Locke’s (1986) book, suggested that “the lab versus field controversy is a false issue and one that scientists and professionals should have put aside long ago” (Campbell, 1986, p. 271). Locke (1986), however, responding to this statement, commented that the dispute “cannot be dealt with adequately at the deductive level. That is why I initiated this book: to examine the issue inductively” (p. 255). This inductive procedure was outlined in the first chapter of the book by Locke (1986) and surrounds the need, through exploratory research, to identify the essential features of field settings that also exist in the lab. The discovery of these essential features, according to Locke (1986), requires an inductive process, as they “cannot necessarily be known in advance” (p. 7).

Another explanation for the lack of research activity may relate to the difficulties associated with compressing multiple-setting studies into the one submission; particularly if within and between-setting analyses are undertaken and discussed. Multiple-setting studies, however, are often encouraged by journals. The incoming editor of OBHDP, for example,

\textsuperscript{1} Although this study was extensive in terms of its depth, with the years 1987-2011 covered, it was less so in terms of its breadth. With only five journals examined, many potentially relevant articles have not been covered in this study, which collectively, could alter this study’s “lack of evidence” finding.
recently wrote “we also encourage investigations that involve a combination of lab experiments and field studies” (Chen, 2011, p. 1).

The field constraints discussed by Ilgen (1986) in his chapter in Locke’s (1986) book may also play a role. Ilgen (1986) suggested that, for various reasons, it is impractical or impossible to conduct field research (e.g., due to constraints such as time, costs, ethics, and threats to health and safety). It could be that much of the lab research to date was conducted for these reasons, with replications not possible. Judging by the fact that most call for their research to be replicated, this does not appear likely. Finally, it may be due to factors underpinned by theories of cognition and motivation, such as cognitive dissonance theory. It might not be in researchers’ best interests, for example, to do twice the work for the one outcome (e.g., one published article). Nor might it be of interest to researchers to replicate their work in other settings. Research appears to be needed to explore if any of the above explanations holds true.

The case for drawing a conclusion in this study that identifies a continuing lack of evidence to support the case for generalizability is also based on the results arising from the small, but growing, number of studies that have looked at the issue; most notably the meta-analyses. Although some of these meta-analytic studies found no between-setting differences (e.g., Horwitz & Horwitz, 2007), most did. Further, most did not discover differences on all the variables examined, but rather only some; a finding similar to that reported by Gersick (1989). This appears to suggest that, if it does eventually emerge, Gist et al.’s (1987) evidence may relate to some between-setting aspects more than others. In other words, using Gersick’s (1989) expression, any differences may not be “wholesale.” What these differences may be is perhaps too early to note, due to the lack of research activity. More inductive research in the tradition of Locke (1986) might help to illuminate what some of these differences may be. Some deductive,
theory-driven research will also help. Indeed, research is needed in a number of areas. Prior to suggesting what these may be, some additional comments will first be made, for they help frame some of these suggestions.

Some Observations

Having examined over 300 team-related articles as part of this study, including those placed in the “other” category, some observations have been made in relation to the issue of generalizability in the teamwork literature. Each of these observations will be presented below. These will then be followed by the research suggestions.

Observation 1: The frequent calls to replicate. Based on the research conducted as part of this study, laboratory to field generalizability appears an important issue in the teamwork literature. The sheer volume of authors raising the issue provides the evidence to support this claim. Although big in terms of the volume of mentions, the issue is not big in terms of research activity. Of the hundreds of generalizability comments examined as part of this study, almost all authors stressed the need for their lab setting research to be replicated in the workplace. This research, however, has yet to be carried out; at least in the five journals examined in this study.

Observation 2: The use of intact teams in laboratory settings. In the team-related literature, the main area of interest is the organizational team, in its many forms. Teams, in most instances, comprise adults, who work together in exchange for pay. They do not usually perform tasks together in exchange for course credit or small monetary incentives. It is therefore surprising to find so many laboratory studies using students as participants, rather than full-time
workers. Further, considering that members of organizational teams are usually interdependent, it is surprising to find so few studies researching intact teams in laboratory settings (i.e., using experimental tasks or conditions). Close inspection of the laboratory studies examined as part of this study found just a few cases where researchers have used intact teams (e.g., Jehn, Rispens, & Thatcher, 2010). Although these studies still have limitations (e.g., the loss of various intrateam interdependencies), they appear underrepresented in the teamwork literature.

**Observation 3: The use of student project groups.** This research discovered that many of the studies conducted by team researchers in the articles examined comprised students. This has frequently involved the use of students operating together as ad hoc groups in laboratory settings, who work on contrived tasks over short time periods. However, it has also involved the use of students performing cooperative or collaborative learning tasks in the physical setting of the classroom. These tasks include experiential learning exercises, business simulations, and complex projects requiring several months to complete, just to name a few. These tasks are also usually performed by the students as part of a course requirement, with grades tied to the group’s performance. It is the use of this latter group type, the student project group, that is of interest here, and for two reasons.

First, those studies using student project groups often appear to imply that these groups are a closer approximate to work teams than laboratory groups, and thus their findings offer stronger generalizability. This often stems from student project groups existing outside of the research context. That is, they are naturally occurring groups, rather than ones merely formed for research purposes. It also stems from the longer periods of time that the project groups exist before disbandment (often weeks, rather than hours), and the tasks performed, which are argued to be more engaging, and more meaningful with “important outcomes” at stake; outcomes such
as grades, and the positive benefits arising from these grades, such as career options (Polzer, Milton, & Swann, 2002). A good example of researchers stressing the strength of their student group sample comes from Kristoff-Brown and Stevens (2001), who argue:

Unlike most previous research on goals in team settings, which has used laboratory experiments with brief interactions among team members performing straightforward tasks (e.g., Tinker-Toy or block-building exercises), this study used a sample of MBA students working on a complex, extended task with significant personal consequences. Although based in an academic environment, our MBA project teams share many characteristics with project teams in organizations. Participants had a personal stake in their teams’ processes and outcomes because these influenced their time available for other classes, course grades, and future recommendations by the instructor. Moreover, despite the fact that these were course activities, the tasks that respondents performed (i.e., diagnosing organizational problems, generating multiple alternatives, advocating a solution) have strong external validity (pp. 1093-4).

Second, student project groups might be problematic in relation to the lab versus field generalizability issue. On one hand, studies using such groups are a type of field study. On the other hand, however, they are not workplace samples, as they involve students operating in a nonworkplace context. Nor is the classroom always a laboratory setting. Sometimes the classroom is transformed into a laboratory setting when students undertake experimental tasks for research purposes. But often, the students in these classroom settings perform nonexperimental tasks designed for learning, and are merely surveyed based on various perceptions. Such research based on students operating in classrooms is therefore a difficult type
to classify. In this study they were treated as “other” and separated from lab studies. This was not the case with the meta-analyses reviewed earlier in this paper; they were treated as laboratory in most instances. The extent to which this has confounded any results is not known. It seems that in the area of team research, the issue of generalizability may be multifaceted, involving various distinctions, such as settings (e.g., laboratory and classroom vs. workplace) as well as samples (e.g., students vs. employees).

**Observation 4: Limits to generalizability and the many differences suggested.** In most instances, those team researchers using lab samples suggest that their study’s findings generalize to work teams. These generalizability statements, however, are often also couched quite cautiously. Part of the reason for this appears to be based on the lack of any strong supporting evidence. Another reason appears to stem from the acknowledgement that the contexts in which their lab samples operate are usually different to those existing in the workplace, and these contextual differences may limit the extent to which their findings are externally valid. Earley and Mosakowski (2000) and Gersick (1989) both noted various contextual differences in their discussions on generalizability. In fact, almost every author’s comments related to generalizability make reference to the question of contextual difference, which is noteworthy for three reasons. First, viewed collectively, there are many contextual differences being proposed. Second, many of these differences are clearly based on careful consideration by experienced team researchers. Third, and perhaps most important, it could offer the blueprint for a possible research agenda, with the first set of researchable contextual factors having already been identified. Just some of the many factors mentioned in the literature are task motivation,
incentives, routinization of tasks over time, organizational culture, member expertise, status and formally appointed leadership.

**Future Research Directions**

In due course, one could envisage a comparative study similar to those contained in Locke’s (1986) edited book, or the one conducted by Anderson et al. (1999). To reach this point, however, a sufficient number of comparative data points would be needed, and this requires more comparative research. It is therefore recommended that a number of between-setting comparative studies be conducted. On one hand, conducting this research will be a large undertaking. On the other hand, though, half of the data already exists. If just a small proportion of the already published studies were to be replicated in the opposite direction (if possible), an extraordinary amount of comparative data could immediately be generated.

Addressing one of the observations presented earlier, there is a need to bring more intact teams into the laboratory. Such research seems to be an important next step for many of the laboratory studies that have used students as subjects. Having discovered that the hypothesised effect “can happen”—the argument often used for conducting laboratory studies—the next step is to discover if the same, or similar effects, *do happen* when the participants are members of intact teams, complete with an interaction history, knowledge of one another’s expertise, status characteristics, and the understanding that they will work again with one another. This will not be an easy task, and may involve the laboratory being taken to the team. It might also require the collaboration with training consultancy firms, who often have access to intact teams for training purposes. Senior management commitment to such research will also be crucial.

Addressing another of the observations discussed earlier, attention needs to be directed towards examining the claims that student project groups offer stronger generalizations to work
teams, than do laboratory groups. A possible way to examine this is to replicate twice, rather than once. That is, to conduct research in workplace, classroom, and laboratory settings, and to compare the results. Research comparing the results deriving from laboratory and classroom settings may also help address this issue, particularly if at least some of the same participants are used, preferably working in the same groups.

It is also recommended that some new multi-setting/sample comparative research be undertaken and that some of the insightful differences mentioned above under Observation 4 be considered for investigation. Theory must also play a part, particularly when it focuses on contextual differences like the notion of “person in situation” (Mischel, 1977; Monson, Hesley & Chernick, 1982) and trait activation theory (Tett & Burnett, 2003; Tett & Guterman, 2000).

To conclude, a number of possible research directions are proposed here. They are offered to address the issue surrounding laboratory to field generalizability with reference to organizational teams. By means of a detailed examination of the teamwork literature, this was found to be an issue that was often mentioned but seldom researched. It was also found to be a researchable one, with many directions possible. It is hoped that these suggestions are considered by team researchers in the future, with the aim being to accumulate more comparative data. By doing so, this may help generate a better understanding of organizational teams. It may also, in time, provide the evidence that Gist et al. (1987) were unable to locate to their satisfaction more than 25 years ago, and that, based on this study, is still hard to locate.
REFERENCES


Bettenhausen, K. L. 1991. Five years of group research: What we have learned and what needs to be addressed. *Journal of Management*, 17: 345-381.


# TABLE 1

## Summary of Meta-Analytic Studies examining Lab versus Field Moderation

| Study                  | Relationships Examined                                                                 | Findings noted by Author(s) and Comments                                                                                                                                                                                                                                                                                                                                                     |
|------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Bell (2007)            | Deep-level composition variables and team performance. Composition variables were: Conscientiousness, agreeableness, extraversion, emotional stability, open to experience, collectivism, preference for teamwork, general mental ability.                                                                                                                                         | Finding noted: “Study setting was a strong moderator of the team composition variable and team performance relationships. In general, consistent with Hypothesis 11, the relationships between personality factors and team performance and between values and team performance were stronger in field settings compared with lab settings” (p. 603). Comment: There were instances where a stronger effect was found in the lab than in the field. For example, the relationship discovered between general mental ability and team performance. |
| Bell et al. (2011)     | Demographic diversity and team performance. Demographic variables were: Race, sex and age.                                                                                                                                                                                                                                                                                                                                                                         | Finding noted: “Study setting moderated the relationships between team performance and race, sex, and age” (p. 728). Comment: Race and sex relationships were stronger in field studies.                                                                                                                                                                                                                                 |
| DeChurch & Mesmer-Magnus (2010) | Team cognition and team effectiveness. Team cognition measures were: Compositional and compilational emergence. Team effectiveness measures were: Process related and performance related.                                                                                                                                                                                                                           | Finding noted: “Setting ...[was]... found to moderate the relationship between compilational cognition and team process such that the cognition–process relationship was stronger in field than in laboratory studies” (pp. 43-44). Comment: However, setting type did not moderate the compositional cognition-process relationship. Nor did it moderate cognition-performance relationships examined (compositional and compilational). |

*table continues*
### TABLE 1
Continued

<table>
<thead>
<tr>
<th>Study</th>
<th>Relationships Examined</th>
<th>Findings noted by Author(s) and Comments</th>
</tr>
</thead>
</table>
| Horwitz & Horwitz (2007) | Team diversity and team outcomes.  
Team diversity measures were: Bio-demographic and task-related.  
Team outcomes were: Quality of performance and social integration. | Finding noted: “Neither criterion measure type nor study setting moderated the relationship between bio-demographic diversity and the quality of team performance” (pp. 1002-3).  
Finding noted: “For the study setting moderator, 37 correlations were analyzed and results indicated that there was no moderating role of study setting in the relationship between team diversity and social integration” (p. 1004). |
| Kleingeld et al. (2011)  | Goal setting and group performance.  
Goal setting measure varied, including: Goal difficulty (easy, moderate, difficult), task interdependence (independent, dependent), and task complexity (low, moderate, high). | Finding noted: “The laboratory–field distinction did not moderate the specific group goal effect ... or the specific difficult group goal effect” (p. 1293).                                                                                           |
| Thatcher & Patel (2011)  | Various variables related to demographic faultlines.  
Relationships included: Age diversity and demographic faultline strength; Demographic faultline strength and team performance; Task conflict and team performance. | Finding noted: “Our results indicate statistically significant differences on many correlations for lab versus field settings.... The strength of the effects for lab studies was consistently stronger than that of field studies” (p. 1127).  
Finding noted: “To further test the moderating effects ... we split our sample into two respective groups to conduct a stacked MA-SEM approach. The differences in paths and the chi-square difference test results are shown in Table 6. Significant path differences were evident for the moderators of study setting (lab vs. field)” (p. 1127).  
Comment: Although the observed effect sizes were consistently stronger, the direction of various relationships differed. The correlation table presenting the difference in Z values for lab v. field (presented as lab minus field) showed both positive and negative relationship, many of which were statistically significant. |
Dear Adam Morgan:

Your paper, "That Laboratory-Derived Findings Generalize to Work Teams: A Search for the Supporting Evidence", with the identification number 14495 submitted to the 2013 Academy of Management Meeting, August 9-13, in Lake Buena Vista (Orlando), Florida has been received. All of the requirements have been met and this submission will be sent for peer review.

Thank you very much for submitting this paper for consideration for the conference program. You will receive notification of the status of your paper in March. If you have any questions, please do not hesitate to contact me.

Now that you have submitted a paper, please sign up as a reviewer and encourage your co-authors and colleagues to do the same. Please visit http://review.aomonline.org/, click the "Sign Up Now" button, and select the OB Division.

I look forward to seeing you at the 2013 AOM meeting in Lake Buena Vista (Orlando)!

Sincerely,

Jacqueline Coyle-Shapiro
OB Division Program Chair
Dear Adam Morgan:

I am pleased to inform you that your paper, submission #14495 titled "That Laboratory-Derived Findings Generalize to Work Teams: A Search for the Supporting Evidence", has been accepted for presentation in a Divisional Paper session at the 2013 Academy of Management Meeting, August 9-13, in Lake Buena Vista (Orlando), Florida. The schedule of sessions is still being finalized. I will send you those details via email when they become available.

Divisional Paper sessions consist of papers with a common theme. Each author will have a set amount of time to present their work. Discussions and a Q&A session will follow after all presentations have been made. A session chair is assigned to run the session and lead the discussion.

To ensure that we have the most accurate information for the program, you will be invited via email to review your personal and submission information. YOU ARE RESPONSIBLE FOR CHECKING YOUR ENTRY. If you have any changes, please email me.

One of the conditions for having your paper accepted is your availability for the entire scholarly program from Sunday to Tuesday, August 11-13, 2013. No changes will be made in the program schedule to accommodate your travel plans.

If you are not able to attend, then you will have to either withdraw the paper or have a co-author present.

You can see the reviewers’ comments for your paper by logging into the submission system, http://submission.aomonline.org/. I hope these comments will help you refine your paper. However, the submission that was reviewed and accepted is what must be presented at the conference. You will also be invited to post your presentation to the online program. More information and instructions on how to do this will be sent in April. Please note that the presentation will be supplementary to the full length papers that will also be posted to the online program. If you decide to post the full length paper, the slides, or both, they will only be available to meeting registrants from May - October.

I would encourage you to complete the rating of reviewers as this information is used to select reviewers for the outstanding reviewer award.

Finally, please remember in order to attend the Annual Meeting and participate on the program in any capacity (eg. presenter, author, discussant, etc.) you must be a member of the Academy and registered to attend the meeting. You can register now at http://aom.org/annualmeeting/registration/register/ in order to take advantage of the early rates.

Again, congratulations on having your paper selected for the 2013 Academy of Management Meeting in Lake Buena Vista (Orlando), Florida. Please contact me if you have any questions.

Sincerely,

Jacqueline Coyle-Shapiro
OB Division Program Chair
Adam Morgan

From: Jacqueline A-M. Coyle-Shapiro <erob.obdivision@lse.ac.uk>
Sent: Thursday, 25 April 2013 1:10 PM
To: Adam Morgan
Subject: 2013 Academy of Management Annual Meeting Paper Status

Dear Adam:

Again, I am pleased to congratulate you for having your paper, submission #14495 titled "That Laboratory-Derived Findings Generalize to Work Teams: A Search for the Supporting Evidence", accepted for presentation in a Divisional Paper session at the 2013 Academy of Management Meeting, August 9-13, in Lake Buena Vista (Orlando), Florida. Please carefully read this entire letter as it includes the following important information.

1. Scheduling details
2. Preparing for the session
3. Registering to attend

SCHEDULING DETAILS

Your paper is included in a session titled Team Processes and Performance. This session is tentatively scheduled on Aug 13 2013 from 1:15PM to 2:45PM in the Europe 6 room of the WDW Dolphin Resort. All presenters in this session are expected to appear at this time. One of the conditions for having your paper accepted is your availability for the entire scholarly program from Sunday to Tuesday, August 11-13, 2013. No changes will be made in the program schedule to accommodate your travel plans. If you are not able to attend, then you will have to either withdraw the paper or have a co-author present.

* NOTE: Scheduled times and locations are subject to change. Please refer to the final printed program or searchable online program for finalized details.

PREPARING FOR THE SESSION

Prior to the Session:
If you have not done so, you can read the reviewers' comments for your submission by logging into the submission system, http://submission.aononline.org/. We hope these comments will help you prepare the presentation. Also, please take a few moments and provide feedback to the reviewers.

ALL accepted papers will be made available online for only meeting registrants to view and download. We hope that this will lead to more interaction and discussion during the session. You will have the option to remove your paper from the online program if you do not want to participate. More information and instructions will be sent regarding this. In the meantime, please still send copies of your paper to the chair of your session by June 1 so that he/she can prepare for the session.

As an additional service you will be able to provide supplements to your paper. You will be able to provide a link to an online posting of a slide show, a PowerPoint presentation, or the audio or video recording of the presentation.
More information and instructions will be sent regarding this.

For the session:
Please prepare a PowerPoint presentation AND BRING A LAPTOP. Each meeting room will be equipped with an LCD projector and screen. Think carefully when preparing your presentation. In particular, I encourage you to visit the 2013 meeting website at http://aom.org/annualmeeting/presenterguide/ and read the materials there, in particular the "Oral Presentation Guidelines". We want to encourage individual paper presentations and entire sessions to be
LAKE BUENA VISTA
ORLANDO
2013

TO INSPIRE AND ENABLE A better NOWLD
Who Am I? Who Are We? The Coevolution of Professional Identity and the Identity of a Profession | Chad Murphy, Pennsylvania State U.

1522 (MSR) Empathy, Consciousness, and Transcendence
1:30pm - 2:45pm WDW Dolphin Resort: Europe 10
Tweet this session: #AOM2013 1522
Organizers: Kathryn Pavlovich, U. of Walkato; Keiko Krahke, U. of Northern Colorado
Ethical Decision Making in Organizations: The Role of Empathy | Emmanuella Patricia Kleingatug, U. of Lausanne; Joerg Dietz, U. of Lausanne
The ACES Decision-Making Technique as a Reframing Tool for Increasing Empathy | Larry Pate, Decision Systems International; Traci Shoblon, Decision Systems International
The Source of Empathy in our Lives: An Explanatory Journey into the Realm of Spirituality | Dana Harajli, Lebanese American University
Developing the Capacity for Managing with Empathy | Sabita Sawhney, Maharishi U.
Empathy and the Brain | Fred Travis, Maharishi U.
Transcendent Empathy: Empathy as the Ability to See the Larger System | Keiko Krahke, U. of Northern Colorado

1523 (Paper Session) - (OB) Interpersonal Processes in Teams: Humility, Helping and More
1:30pm - 2:45pm WDW Dolphin Resort: Europe 11
Tweet this session: #AOM2013 1523
Chair: Stephen E. Humphrey, Pennsylvania State U.
Leading by Modeling: From Leader Prosocial Motivation to Team Effectiveness | Jia (Jasmine) Hu, U. of Notre Dame; Robert C. Liden, U. of Illinois, Chicago
How Perceptions of Deservingness and Permanence Affect Peripheral Group Members’ Helping Behavior | Tina R Opie, Babson College
Humility in Teams: Collective Humility and its Impact on Team Growth, Climate, and Performance | Bradley O. Owens, SUNY Buffalo; David R. Hekman, U. of Colorado
The Contextualized Self: How Team-Member Exchange Leads to Coworker Identification and Helping OCB | Steven M. Farmer, Wichita State U.; Linn Van Dyne, Michigan State U.; Dihan Kamarl, Indiana University of Business

1524 (Paper Session) - (OB) Team Processes and Performance
1:30pm - 2:45pm WDW Dolphin Resort: Europe 6
Tweet this session: #AOM2013 1524
Chair: Sean Walker, U. of Tennessee at Martin

- Examining the Asymmetrical Effects of Goal Failures in Groups | Ke Mai, U. of Arizona; Aleksandr P. J. Ellis, U. of Arizona; Jessica Siegel Christian, U. of North Carolina
- The Influence of Representational Gaps on Team Processes and Performance | Shirley Chaoyi Wang, Cornell Johnson School; Elizabeth A. Mannix, Cornell U.
- Productive Dialogue and Creativity Performance | Yun-Hwa Chiang, Ming-Chuan U.; Chu-Chun Hsu, Southern Taiwan U. of Science and Technology; Hai-An Shih, National Cheng Kung U.; Wen-Yen Hsu, National Cheng Kung U.
- That Laboratory-Derived Findings Generalize to Work Teams: A Search for the Supporting Evidence | Adam Charles Morgan, U. of Technology, Sydney
- Why Boundary Spanners Perform Well? | Yanan Wang, Renmin U. of China; Zhen Wang, Renmin U. of China

1525 (OB) Fostering Coworker Relationships: Let’s Share Knowledge and Food!
1:45pm - 2:45pm WDW Dolphin Resort: Europe 7
Tweet this session: #AOM2013 1525
Chair: Anders Dyasvik, BI Norwegian Business School
Coworker satisfaction in 3D: Perceiver, Partner, and Relational Coworker Satisfaction | Elia Bybington, Australian School of Business, UNSW
Cooperating Over Food and Firefighting: Organizational Benefits of Commensality | Kevin M. Kniffen, Cornell U.; Brian Wansink, Cornell U.; Carol M. Devine, Cornell U.; Jeffery Sobal, Cornell U.
Sources of Coworker Relationships: Social Relations Modeling of Relational Models | Elia Bybington, Australian School of Business, UNSW
When does Coworker Knowledge Sharing Matter to Performance? | Seokyoung Loretta Kim, Seoul National U.; Seoju Han, Seoul National U.; Chang Won Go, Seoul National U.; Seokhwa Yun, Seoul National U.

1526 (OB) Witnessing and Experiencing Abusive Supervision
1:50pm - 2:45pm WDW Dolphin Resort: Ocean Rivera
Tweet this session: #AOM2013 1526
Chair: Cynthia Kay Stevens, U. of Maryland
Moderation of Career Aspiration in the Relationship between Abusive Leadership and Work Outcomes | Jie Shen, U. of South Australia; Ningyu Tang, Shanghai Jiao Tong U.
Roles of Gender and Identification on Abusive Supervision and Proactive Behavior | Kuo Yung, Hong Kong Polytechnical U.; Wing Lam, Hong Kong Polytechnical U.; Ziguang Chen, City U. of Hong Kong
Power-Dependency Dynamics and Abusive Supervision | Elijah Xun Ming Wei, U. of Maryland; Hui Liao, U. of Maryland
Most Innovative Student Paper sponsored by Sage Publications (a student must be the sole or first-author and must have taken the lead role on the project)
Supervisor Support: Does it Buffer or Exacerbate the Adverse Effects of Supervisor Undermining? | Inbal Bille Nahum-Shani, U. of Michigan, Ann Arbor; Sandy Lim, National