Group Decision and Negotiation **15:** 389–406, 2006.
DOI: 10.1007/s10726-006-9042-x


# The Effects of Cultural Diversity in Virtual Teams Versus Face-to-Face Teams

D. SANDY STAPLES AND LINA ZHAO
*Queen's School of Business, Queen's University, Kingston, Ontario, Canada K7M 6K5*
*(E-mail: sstaples@business.queensu.ca)*

***Abstract***

Diversity in the national background and culture of team members is common in virtual teams. An experimental study, with short term teams, was undertaken to examine the effect of cultural diversity on team effectiveness and to examine if this effect changes depending if the team worked face-to-face (F2F) or virtually. Heterogeneous teams were created that had greater diversity than homogeneous teams of individualism/collectivism values, different languages spoken, country of birth, and nationality. The teams worked on a desert survival task either F2F or virtually (via audioconference and electronic chat tools). The overall results indicated that heterogeneous teams were less satisfied and cohesive and had more conflict than the homogeneous teams, although there were no statistical differences in team performance levels. However, examining just the heterogeneous teams found that the performance of the virtual heterogeneous teams was superior to that of the F2F heterogeneous teams. The results support Carte and Chidambaram's (2004) theory that the reductive capabilities of collaborative technologies are beneficial for newly-formed diverse teams.

**Key words:** team diversity, virtual teams, cultural diversity, national culture, experiment, survival task, reductive capabilities of collaborative technologies

## Introduction

Teams offer companies potentially effective ways to combine the various skills, talents and perspectives of a group of individuals to achieve corporate goals. With globalization, transnational teams are a reality (Earley and Mosakowski 2000) and need to be studied (Hambrick et al. 1998). These teams are usually made up of members from different national backgrounds, meaning the members come from different national cultures, possibly speak different languages, and were raised in different countries that may have different value systems. Today's electronic communication capabilities makes it easier and common for these team members to work together while being physically located in different cities or countries. These geographically-distributed teams are commonly referred to as virtual teams. The degree of geographic dispersion within a virtual team can vary widely from having one member located in a different location than the rest of the team to having each member located in a different country. Virtual teams that span multiple countries create the strong possibility that members of the team will have diverse national backgrounds (Evaristo 2003; Powell et al. 2004). Understanding the possible impacts of this diversity on team performance is important for today's organizations.

The purpose of this study is to examine the effect of cultural diversity on team processes and performance and to examine if this effect changes depending on the communication mode used. In their theory paper, Carte and Chidambaram (2004) proposed that the reductive capabilities of collaborative technologies (e.g., electronic tools such as email, group support systems, computer conferencing) can reduce the negative effects of diversity early in the life of a diverse team. If the effect of cultural diversity is different for teams communicating electronically versus those communicating face-to-face (F2F), this may have important implications for the design of virtual teams (i.e. those using information and communication technologies (ICT) as their primary communication media) and their organizations. For example, as Carte and Chidambaram suggested, diverse virtual teams may be better off not meeting F2F until relationships have been developed. Minimizing the salience of surface-level diversity by avoiding F2F meetings early in the life of team may reduce the potential negative impact of this diversity.

An experimental method was used to examine the effect of cultural diversity on F2F and virtual short-term, project teams. In this experiment, virtual teams communicated via audioconference (telephone conferencing) and a synchronous chat system. An audioconference phone system was chosen as the main communication tool for the virtual teams since studies of several on-going industry-based virtual teams (Staples et al. 2004) found that the most common way teams met were via audioconferencing, with other electronic tools used as needed during the audioconference meeting (e.g., electronic white-boards, synchronous chats, instant messaging). Choosing audioconference and synchronous chat as the communication channels for the virtual teams also addresses Baltes et al. (2002) call for more relevant research using industry-adopted communication tools, such as audioconferencing.

This paper is organized as follows. In the next section, the theoretical background and relevant literature are presented and the hypotheses for the study are developed. The methodology used to test the hypotheses is then presented, followed by a description of the results. The final section discusses the findings and limitations, and offers suggestions for future research.

## Theoretical Background and Development of Hypotheses

As reviewed by Milliken and Martins (1996), different aspects of diversity in teams have been investigated including observable differences such as race, ethnicity, gender, and age, and unobservable differences such as skills, information and knowledge, values, cognitive processes, and experience. Variety in overt characteristics (readily observable traits) is typically referred to as surface-level diversity, while variety in unobservable characteristics is referred to as deep-level diversity (Carte and Chidambaram 2004).

There can be both positive and negative aspects of team diversity. Value in diversity comes from increased creativity, innovation and flexibility (Jehn, Northcraft and Neale 1999; Lau and Murninghan 1998; McLeod et al. 1996). A variety of perspectives and experiences (i.e., deep-level diversity) can bring more information and ideas into the team, stimulate thinking, and can bring different networks of contacts and resources to the team. When the deep-level diversity is relevant to the task facing the team, higher-quality outcomes should result; however, if the variety is not relevant to the task, then there is no basis for expecting

the diversity to enhance the team's activity. Therefore, the beneficial effect of deep-level diversity within a team is contingent on the relevance of the diversity to the team's task (McLeod et al. 1996).

Negative aspects of team diversity include communication difficulties, misunderstandings, decreased cohesion and increased conflict. These process losses result in decreased performance and satisfaction (Hambrick et al. 1998; Lau and Murninghan 1998; Williams and O'Reilly 1998). Social identity theory, social categorization theory and the similarity/attraction paradigm suggests that the negative effects associated with diversity are due to the creation of in-groups and out-groups (Carte and Chidambaram 2004; Salk and Brannen 2000). People implicitly categorize themselves into subgroups according to salient cues and identify more closely with people they perceive as being similar to themselves. They do this to achieve and maintain positive self-identity. As in- and out-group characteristics become salient within subgroups, individuals become more biased towards their subgroup. Emotional attachments to the subgroups become potential sources for interpersonal and relationship conflict with members of other subgroups. Relationship conflict reduces satisfaction and team performance. Lower personal attraction to group members, higher turnover, and poorer communication also results in reduced social integration and cohesion. If group members are diverse on multiple attributes that align together, strong faultlines can develop which create further subgroup problems and team development problems (Lau and Murninghan 1998).

The effect of the different types of diversity on team outcomes is significantly affected by time (Carte and Chidambaram 2004). When groups newly form, they have little information about each other than observable characteristics. Members try to make sense of each other and their task. If surface-level diversity is present, members may use salient characteristics to implicitly categorize themselves into subgroups, creating the negative effects described above (Lau and Murninghan 1998). Although surface-level traits are immediately apparent upon team formation, deep-level traits become salient as the team members interact over time. The potentially positive effects of deep-level diversity take time to emerge. Therefore, in the short-term, diverse teams typically perform worse than homogeneous teams due to the early impacts of surface-level diversity. Empirical research has found this pattern in that team diversity benefits seem to be obtained after a team has interacted for some time and relational ties have been developed (Carte and Chidambaram 2004; McLeod et al. 1996; Watson et al. 1993).

While there are many different types of diversity, this study focuses on differences in the national and cultural background of team members. Cultural background has elements of both surface-level characteristics and deep-level characteristics. Surface level characteristics that can vary depending on where one was born and/or nationality include race and ethnic characteristics and native language. Deep-level characteristics that will vary for people from different countries are cultural values. The national culture one grows up and lives in influences thinking, expectations, and behavior (Evaristo 2003; Hambrick et al. 1998). This study focuses on surface-level cultural diversity specifically in terms of nationality, country of birth and native language, and on one deep-level aspect of cultural diversity – individualism/ collectivism values. This aspect of national culture and the reason for including it is explained next.

National culture is a complex construct to define and study. A common definition of national culture (Evaristo 2003) is offered by Hofstede (1980). He defines culture "as a collective phenomenon, because it is at least partly shared with people who live or lived within the same social environment where it was learned. It is the collective programming of the mind that distinguishes the members of one group or category of people from another." Hofstede's research defines national culture in terms of five value dimensions: individualism/collectivism (personal interests versus group interests), power distance (acceptance of inequality), uncertainty avoidance (dislike for ambiguity), masculinity/femininity (assertiveness and focus on work goals versus personal and family goals), and time horizon (sometimes called Confucian dynamism – short-term versus long-term orientation) (for more information on the dimensions and how they were derived, see Anderson and Hiltz 2001, Bond 1988, Evaristo 2003 or Hofstede 1983).

While recognizing the importance of the five dimensions, this study focuses on the individualism/collectivism dimension due to its potential relevance to the phenomenon being study – team performance. Individualism/collectivism indicates how an individual puts his/her interests ahead of the team's interests. Research has found that people from collectivist cultural backgrounds are more willing to help people, make personal sacrifices and are more cooperative than people from individualist cultural backgrounds (McLeod et al. 1996). Therefore, individualism values potentially affect communication and coordination patterns among individuals working in teams and their expectations (Earley 1989). In high individualistic cultures, people rely on the use of words to convey meaning whereas in low individualistic/high collectivistic cultures, tone of voice, timing, facial expressions and behavior are also important parts of the communication (Anderson and Hiltz 2001). The individualism/collectivism dimension has also been used in previous research to represent respondents' national culture (e.g., Kessapidou and Varsakelis 2002; Stedham and Yamamura 2004), and Sondergaard (1994) suggests it is the most validated of Hofestede's five dimensions.

It is expected that due to the in-group and out-group problems described above, surface-level diversity regarding national background will lead to lower cohesion and greater conflict in the early life of a team. This in turn leads to lower team performance and satisfaction. Teams that are highly diverse on individualism/collectivism values will have people on the team that have different expectations and values regarding communications and interaction patterns. Some people will be willing to cooperate and sacrifice for the overall team, and some will not. This also could create conflict and lower cohesion, contributing to lower team outcomes. Differences in native languages can also contribute to communication difficulties within teams (McDonough et al. 1999). If the cultural diversity is not beneficial to the task the team is performing, then there would be no offsetting increase in performance over time and the diverse teams would be expected to perform worse than homogeneous teams. Thus, in this context, our first hypothesis is:

Hypothesis 1 – Higher levels of cultural diversity will be associated with lower cohesion and higher conflict, and lower team outcomes (i.e., team performance and satisfaction).

Carte and Chidambaram (2004) propose that communication technologies have bundles of capabilities and these capabilities can be categorized as two types: reductive or additive. Reductive capabilities reduce aspects of communication and speech patterns that would be present in traditional face-to-face communication. Reductive capabilities include visual anonymity (identification is limited), equality of participation (normal turn taking may be reduced), and asynchronous communication (immediate feedback is limited). Additive capabilities enhance normal communication exchanges and include coordination support (tracking resources and project progress), electronic trails (creating records and retrieving information) and enhanced functions (decision making tools, file transfers and rich messaging).

Carte and Chidambaram (2004) further suggest that the bundles of capabilities are most useful at different stages of a diverse team's development. Specifically, reductive capabilities are valuable early in the life of a diverse team, whereas the additive capabilities will add value later in a team's life (i.e., after a shared team identity is established), by providing support for decision-making and coordination. This study is particularly focused on the potential benefits of reductive capabilities in the functioning of a newly-formed diverse team. The key reductive capability is visual anonymity since this reduces the immediate saliency of surface-level diversity. Team members can assess inputs from others and form opinions based on merit, without surface-level diversity affecting judgments. Decreasing perceived surface-level diversity potentially reduces the team member's categorization processes, thereby reducing the formation of perceived in-group and out-groups. Less disintegration of the team into subgroups should improve team interaction processes by reducing interpersonal disagreements and conflict, resulting in higher cohesion, team performance and team satisfaction. Enhanced equality of participation can increase participation, allow minority opinions to be heard, and foster a sense of belonging to a group. Asynchronous communication can also benefit a newly-formed team since members will have time to consider how they should say things, potentially reducing quick reactions, miscommunication and attribution errors (Carte and Chidambaram 2004).

As mentioned in the introduction, to mirror virtual team practice, the virtual teams in this study communicated with telephone conference and electronic chat tools, with the participants choosing how frequently to use one or both of the tools. The reductive capabilities of these communication media would not be at the extreme ends of the potential to reduce salience of diversity (Carte and Chidambaram 2004); however, some reductive capabilities are certainly present in this bundle of communication media. For example, although communicating with media richer than a pure text-based system, our virtual teams did not have visual contact and they communicated using disguised names. Therefore, visual anonymity was present. However, telephone use does allow one to hear tone and voice expression (such that the amount of verbal cues and ability to notice different language accents are relatively high compared to another medium like an electronic text-based message system (Baltes et al. 2002)). Telephone conference is high on synchronicity since the conversations happen in real-time and the chat tool used was a synchronous tool. Therefore, the asynchronous communication reductive capability potential benefits were very limited in this study. Equality of participation through the electronic chat tool and the telephone would potentially be higher than face-to-face communication, although not as high as in a pure

text-based chat system. Overall, consistent with Carte and Chidambaram's (2004) propositions, we expect the reductive capabilities in the communication media used by virtual teams to reduce the salience of surface-level diversity. In addition, we also expect that the elimination of visual cues will reduce the visibility of different communication and interaction styles, and the visibility of negative reactions to this diversity, reducing the negative impact on a team having members with different individualistic/collectivistic expectations and values. Hence, we propose that the negative effects of cultural diversity (surface-level diversity and individualism value diversity) in virtual teams would be reduced by the lack of traditional face-to-face communication. In teams with low diversity (i.e., culturally homogeneous), we do not expect to see differences between the face-to-face and virtual teams. Thus:

Hypothesis 2 – Communication mode moderates the relationships between cultural diversity and team processes and outcomes. Specifically, culturally heterogeneous virtual teams using communication technologies with reductive capabilities will have higher cohesiveness, less conflict, better task performance, and higher satisfaction than culturally heterogeneous F2F teams.

The methodology used to test these hypotheses is described next.

## Methodology

### *Participants*

Seventy-nine teams participated in the 2 × 2 experiment (see Table 1). Team size was designed to be 5 people, although some last minute no-shows resulted in 4-person teams (which was deemed to be acceptable in terms of stimulating enough interaction; three person teams were not allowed). In total, 380 people participated. For adequate power, a minimum of 15 teams per cell (Thompson and Coovert 2003) was required and exceeded for all cells.

Participants were university students (60% undergraduate; 40% graduate students) with a fairly even gender split (i.e., 58% female). Diversity of student mix and diversity of gender

*Table 1*. Experiment design and sample information.

| | Cultural diversity | |
|---|---|---|
| | Homogeneous (low) | Heterogeneous (mixed – high) |
| | Communication mode | |
| F2F | 21 teams<br>12 teams∗ 5 person + 9 teams∗ 4 person ($n = 96$) | 19 teams<br>18 teams∗ 5 person<br>+ 1 team∗ 4 person ($n = 94$) |
| Virtual | 20 teams<br>17 teams∗ 5 person<br>+ 3 teams∗ 4 person ($n = 97$) | 19 teams<br>17 teams∗ 5 person<br>+ 2 teams∗ 4 person ($n = 93$) |

*Table 2*. Background of subjects.

| Country of birth | Number of people | Percentage (%) |
|---|---|---|
| Canada or U.S.A. | 195 | 51.3 |
| Asia | 126 | 33.2 |
| Europe | 17 | 4.5 |
| Africa | 11 | 2.9 |
| South America | 11 | 2.9 |
| Middle East | 9 | 2.4 |
| Mexico | 8 | 2.1 |
| Central America | 2 | 0.5 |
| Australia | 1 | 0.3 |
| | 380 | |

was similar in the four cells (i.e., there were no statistical differences on Blau heterogeneity indexes: $F_{(3,75)} = 0.9$, $p > .10$; $F_{(3,75)} = 0.2$, $p > .10$, respectively). The two largest groups (see Table 2) in terms of country of birth were people born in Canada or the U.S.A. (51%) and in Asia (33%), providing good variance on cultural values (Hofstede 1983) and surface-level diversity.

Participation of subjects was voluntary and was not linked to any courses or course credit. Subjects each received $15 for participating. As a team performance incentive, the top teams in each treatment cell were also given a $20 per person bonus. Subjects were drawn from across multiple faculties of a large university campus and were combined into teams with members from different areas as much as possible to minimize previous knowledge of fellow team members. To test for previous history, we did ask subjects in a post-questionnaire if they knew any of their teammates prior to the start of the experiment (i.e. knew them in any way – we didn't ask if they worked with each other before). Seventy percent knew nobody on the team and 24% knew only one person, implying that the team did have very low prior history since 94% knew none or only one member of the team prior to the study.

### *Decision task*

A decision-making task that required interaction and communication and that had an expert solution available (to create a measure of the team's performance) was needed. We also wanted a task where the cultural diversity was unlikely to be of benefit so that the potential positive influences of different views did not confound the possible negative effects of diversity on interactions and team processes. Johnson and Johnson's (1994) desert survival task was chosen since it was a scenario that few people would have relevant experience in and cultural background was not likely to be beneficial. In this task, participants first read a short document that places them into an airplane crash scenario in a desert with the rest of their teammates. There are twelve items that participants have to rank in order of their importance for survival. This is done first individually and then as a team. This task has been used frequently in small group research (e.g., Bottger and Yetton 1987; Haslam et al.

1998; Rogelberg and O'Connor 1998; Straus 1996) and has similarities to problems dealt with by temporary teams on the job (Thompson and Coovert 2003). It is mostly a Quadrant II task in McGrath's (1984) group task circumplex, since it is an interdependent intellective task (Potter and Balthazard 2002). It requires teams to solve a problem that has a correct answer (i.e., an expert's answer). It also has aspects of a judgment/decision making task since a team member can not prove the correctness of his/her answer and has to persuade teammates, and has aspects of a negotiation/cognitive-conflict task since teammates have to discuss and resolve differing opinions regarding survival strategies and the ranking of the items (Thompson and Coovert 2003).

*Design and procedure*

Four experimental conditions corresponding to two media types (i.e., communication mode) and two degrees of cultural diversity were created (see Table 1). The face-to-face (F2F) teams met in a room and worked on the task around a table without any computer tools. The virtual teams did not meet F2F. They worked in separate rooms (within the same building) so they could not see each other as they interacted. They communicated as a team via telephone conference and a synchronous electronic chat system on PC's (implemented through Lotus Notes – see Figure 1). It was up the team's members to decide how they used the media choices. Some teams used both extensively, whereas others tended to use one medium more than the other, with telephone usually being the medium of choice in the situation.

The cultural diversity treatment was created using information gathered from subjects in a pre-questionnaire they completed before they were assigned to experimental teams. This allowed us to collect information used to create homogeneous and heterogeneous groups. Two of the indicators of cultural diversity were used to create these teams. A score for the individualism/collectivism index of Hofstede's Values Survey Module 1994 (VSM 94) was used (using Hofstede's formula), along with information on the subject's first language. Variety of subjects' language has been previously used as an indicator of diversity of national background (e.g., Anderson and Hiltz 2001; Mortensen and Hinds 2001) and the individualism/collectivism dimension has also been used in previous research to represent respondents' national culture (e.g., Kessapidou and Varsakelis 2002; Stedham and Yamamura 2004). Creating teams that had dissimilar individualism/collectivism scores (i.e., the high diversity/heterogeneous teams) placed people on the same team that potentially had different expectations and values regarding communications and interaction patterns, which, as previously explained, could affect team processes and outcomes.

The subjects were assigned into teams of five such that the homogeneous teams had members with all the same first language (i.e., no diversity on first language characteristics) and similar values on their individual scores on the individualism index. Specifically, the team members' scores of the individualism index had a standard deviation of less than twenty (the range obtained for the VSM individualism scores was −45 to 225 so a standard

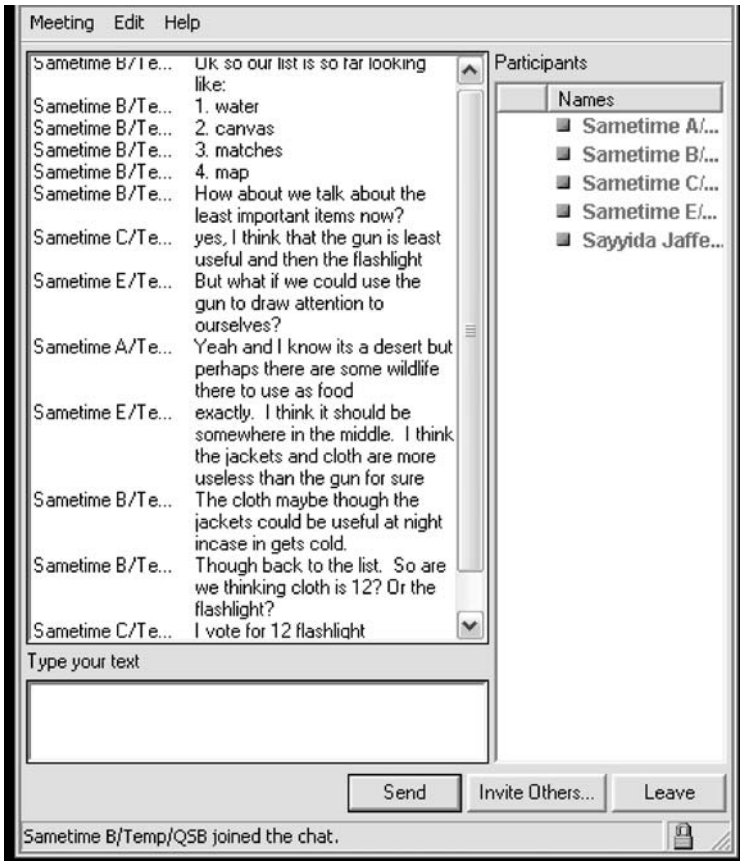


*Figure 1*. A sample screen view of the synchronous chat system.

deviation of 20 indicates fairly similar individual scores). For the culturally heterogeneous teams, there were two or more first languages among the team members (i.e., diversity within the team on first language background) and their scores of the individualism index measure had a standard deviation of more than forty (i.e. much more variance in the scores than in the homogeneous teams). The effectiveness of this procedure was checked upon completion of the experiments. The average standard deviation for the homogeneous teams and heterogeneous teams were 16.6 (standard deviation of 6.7), and 55.5 (standard deviation of 12.7), respectively. These means are statistically significantly different ($t_{77} = 16.8$, $p < .001$), indicating that the creation of the diverse and similar teams on this dimension was successful. Note, this treatment was not attempting to create teams that were high or low on the individualism index; the purpose was to create teams with members that had similar values (i.e. low diversity) and teams with members that had different values (i.e., high diversity). Consistent with this, the average individualism index scores for the homogeneous

and heterogeneous teams were relatively similar (97.2 and 87.6, respectively; $t_{77} = 1.6$, $p > .05$). The mean team values (ranging from 39 to 165) indicate that, on average, the teams held scores closer to the individualism end of the scale (i.e., Hofstede's VSM manual suggests that 100 indicates individualism, whereas scores approaching zero indicates collectivism).

The diversity of team members' reported nationality and country of birth (collected in the pre-questionnaire) were examined by creating Blau's (1977) index of heterogeneity (D) for these two indicators of national culture. Blau's nationality index values for the diverse treatment teams and the homogenous treatment teams were statistically different at 0.53 and 0.16, respectively (($t_{77} = 8.48$, $p < 0.001$). Blau's country of birth index values for the diverse and homogenous teams were also statistically different at 0.63 and 0.21, respectively (($t_{77} = 9.64$, $p < 0.001$). The specific diversity of country of birth, combined into regions (see Table 2), was also examined. For the homogeneous treatment teams, 54% of all teams had all members born in the same region, 37% of the teams had members from two regions, 7% were from three regions, and 2% of the teams had members from four regions. For the heterogeneous treatment teams, no teams had all team members born from the same region, 32% had people born in two of the regions, 50% had people born in three regions, and 18% had people born in four of the regions. This analysis supports that the treatment did create teams with different levels of diversity, with respect to the indicators of national background and culture that were measured.

*Experimental procedure*

Participants first completed a questionnaire that collected information on their background and their cultural values as per the VSM questions (Hofstede 1994). Based on that information, the homogeneous and heterogeneous teams were created and scheduled for the task. Upon arriving, members of the F2F teams were led to one room and asked to sit around a table. The experiment started with participants being given the description of the task description and then being given ten minutes to individually read and rank the survival items. They then worked together for up to forty-five minutes to discuss and agree on a ranking that was submitted as their team's recommendation. Upon arrival, members of the virtual teams were individually given a short training session (less than five minutes since the system was quite simple) for the synchronous chat tool and then they waited in their separate offices until all team members were ready to begin. When they were all ready, team members were also given ten minutes to work individually. At the end of this period, the coordinator phoned all team members to establish the conference call. The electronic chat system was also activated at the same time. The team discussed their views for a maximum of 45 minutes and agreed upon a solution that was submitted to the coordinator electronically. A post-questionnaire was completed by all subjects after their team's solution had been submitted. The post-questionnaire contained measures for the dependent variables, except for the performance construct (which was obtained by comparing the team's answer with the expert's answer, as per the desert task manual).

*Construct measurement*

The pre-questionnaire gathered demographic data and measured Hofstede's VSM five cultural dimensions (Hofstede 1994). The post-questionnaire measured the individual's satisfaction with team process using Green and Taber's (1980) scale (Cronbach's alpha = .77). Cohesion (Cronbach's alpha = 0.79) and conflict (Cronbach's alpha = 0.65) were measured using Lind's (1999) scales, which were based on Van De Ven and Ferry's (1980) measures. Team performance was assessed by comparing the team's recommendation with the expert ranking of the survival items. The unit of analysis was at the team level so the individual responses for satisfaction with team process, cohesion and conflict were aggregated to the team level. Average $r_{wg}$ (James et al. 1984) values were 0.88, 0.85 and 0.78 respectively, indicating it was valid to perform the aggregation (i.e., greater than 0.70). Analysis to determine statistically significant differences between the treatments groups was done with MANOVA analysis (H1) and t-tests (H2 – due to the reduced sample size). One-tail p-values were used, consistent with the directionality of the hypotheses.

## Results

The construct scores for the dependent variables are presented in Tables 3 through 6. Hypothesis 1 was tested with MANOVA procedures. Statistically significant differences between cells E and F were examined for all four dependent variables. The Omnibus test was significant ($F(4,74) = 2.80$; $p = 0.03$). Univariate F-tests showed that the team attitudes and conflict were all statistically significant (Satisfaction with Team Process $F(1,77) = 9.78$; $p = 0.001$; Cohesion $F(1,77) = 7.88$; $p = 0.003$; Conflict $F(1,77) = 3.32$; $p = 0.036$). Team performance was not statistically significant ($F(1,77) = 0.05$; $p = 0.42$). Examination of the construct scores showed that cohesion and satisfaction were higher in the homogeneous teams and conflict was lower. Therefore hypothesis 1 was partially supported in that the diverse teams had poorer team processes (cohesion and conflict) and lower satisfaction; however, the performance of these teams was not statistically less than the homogeneous teams.

Hypothesis 2 was also partially supported. Examination of only the culturally diverse (i.e., heterogeneous) teams, found that there was a statistically significant difference in the team performance (cell D versus B: $t(36) = 1.71, p = 0.048$). The cell means (B and D) show that team performance scores were superior for the virtual heterogeneous teams compared to the F2F heterogeneous teams, indicating the virtual team rankings were closer to the expert ranking (i.e., they performed better). There were no statistically significant differences in the performance of the F2F homogeneous teams and the virtual homogeneous teams. There were no statistically significant differences in any of the team attitudes or conflict levels across the F2F and virtual homogeneous teams or across the F2F and virtual heterogeneous teams.

*Table 3*. Satisfaction with team process scores (a low score indicates a more positive view of team process; aggregated for each team).

| | Diversity | |
|---|---|---|
| | Homogeneous (low) | Heterogeneous (mixed – high) |
| | Communication mode | |
| F2F | Cell A | Cell B |
| | Mean = 1.68 | Mean = 1.88 |
| | SD = 0.29 | SD = 0.39 |
| | $N = 21$ | $N = 19$ |
| Virtual | Cell C | Cell D |
| | Mean = 1.55 | Mean = 1.80 |
| | SD = 0.27 | SD = 0.33 |
| | $N = 20$ | $N = 19$ |
| (Combined) | Cell E | Cell F |
| | Mean = 1.61 | Mean = 1.84 |
| | SD = 0.29 | SD = 0.35 |
| | $N = 41$ | $N = 38$ |

*Table 4*. Cohesion scores (higher scores = more cohesive team; aggregated for each team).

| | Diversity | |
|---|---|---|
| | Homogeneous (low) | Heterogeneous (mixed – high) |
| | Communication mode | |
| F2F | Cell A | Cell B |
| | Mean = 4.39 | Mean = 4.17 |
| | SD = 0.37 | SD = 0.37 |
| | $N = 21$ | $N = 19$ |
| Virtual | Cell C | Cell D |
| | Mean = 4.42 | Mean = 4.20 |
| | SD = 0.26 | SD = 0.37 |
| | $N = 20$ | $N = 19$ |
| (Combined) | Cell E | Cell F |
| | Mean = 4.40 | Mean = 4.19 |
| | SD = 0.32 | SD = 0.36 |
| | $N = 41$ | $N = 38$ |

## Discussion

In a recent review of virtual team research, Powell et al. (2004) suggested that diversity may be less apparent in virtual teams, potentially reducing the process losses caused by cultural heterogeneity. Carte and Chidambaram (2004) went considerably beyond this to suggest how different types of diversity affect team functioning, propose how different types of electronic communication can affect the impact of diversity, and propose how time changes the effects. The main contribution of our paper is to test a subset of Carte and Chidambaram's propositions. Whether or not the reductive capabilities of communication

*Table 5*. Conflict scores (higher = less team conflict: aggregated for each team).

| | Diversity | |
|---|---|---|
| | Homogeneous (low) | Heterogeneous (mixed – high) |
| | Communication mode | |
| F2F | Cell A | Cell B |
| | Mean = 4.35 | Mean = 4.14 |
| | SD = 0.43 | SD = 0.48 |
| | $N = 21$ | $N = 19$ |
| Virtual | Cell C | Cell D |
| | Mean = 4.43 | Mean = 4.30 |
| | SD = 0.35 | SD = 0.34 |
| | $N = 20$ | $N = 19$ |
| (Combined) | Cell E | Cell F |
| | Mean = 4.38 | Mean = 4.22 |
| | SD = 0.39 | SD = 0.42 |
| | $N = 41$ | $N = 38$ |

*Table 6*. Team performance scores (lower score = higher performance in terms of matching experts solution).

| | Diversity | |
|---|---|---|
| | Homogeneous (low) | Heterogeneous (mixed – high) |
| | Communication mode | |
| F2F | Cell A | Cell B |
| | Mean = 42.52 | Mean = 46.00 |
| | SD = 12.07 | SD = 9.12 |
| | $N = 21$ | $N = 19$ |
| Virtual | Cell C | Cell D |
| | Mean = 43.00 | Mean = 40.53 |
| | SD = 9.84 | SD = 10.58 |
| | $N = 20$ | $N = 19$ |
| (Combined) | Cell E | Cell F |
| | Mean = 42.76 | Mean = 43.26 |
| | SD = 10.90 | SD = 10.13 |
| | $N = 41$ | $N = 38$ |

media could reduce the negative effects of national background and cultural diversity in newly-formed teams was examined.

This idea (Hypothesis 2) was partially supported indicating that the reductive capabilities with communicating via telephone and electronic chat did reduce the negative impact of team diversity. Although performance was higher in the virtual heterogeneous team compared to the F2F heterogeneous team, there were no statistical differences in the satisfaction, cohesion or level of conflict (although all had slightly better mean values for the virtual teams). Post-hoc analysis was done to examine the robustness of the findings. The teams were split into those that were high on the Blau's heterogeneity index for nationality (i.e., highly heterogeneous = high diversity) and teams that were low (i.e., homogeneous). In the

heterogeneous teams, the virtual teams had statistically better performance and statistically less conflict than in the heterogeneous F2F teams. There were no significant differences in the homogeneous teams. These findings are consistent with the findings of Hypothesis 2 and add support to the idea that reductive capabilities may also reduce conflict.

These results have important implications for practitioners. Conventional wisdom is that virtual teams should have a project kick-off meeting that is held F2F. This helps establish social bonds and relationships. However, as suggested by Carte and Chidambaram (2004) and supported by the results of this study, this practice should only be followed if the teams are homogeneous. If teams are diverse, especially on surface-level elements, then rich media meetings, such as F2F, should be avoided until a team identity has been established. Teams should communicate using collaborative technologies that have reductive capabilities. In this way, the creation of subgroups that hurt team processes and outcomes is minimized. If the surface-level diversity is low in the team, then the practice of early F2F meeting is likely advantageous.

Partial support was also found for hypothesis 1 in that diverse teams overall (i.e., teams in both communication treatment modes combined) had more negative attitudes toward their team (satisfaction and cohesion) and more conflict. This is consistent with previous theory and research regarding the negative effect of cultural diversity on team interactions (e.g., Johansson et al. 1999; Kayworth and Leidner 2000; Maznevski and Chudoba 2001; Sarker and Sahay 2002; van Ryssen and Godar 2000). Interacting with people with similar cultural values, native language and national background leads to a cohesive team and less conflict presumably because subgroup fractures do not form.

### *Ideas for future research and limitations*

There are many possible extensions for this work. We simply examined two levels of diversity – low and relatively high. Examining different degrees of heterogeneity could be valuable to examine if and when faultlines develop (Lau and Murninghan 1998). Homogeneous groups are similar so faultlines do not develop. In highly heterogeneous teams, few commonalities exist for the basis of creating in-groups and faultlines (Earley and Mosakowski 2000). However, in-groups should be strong under moderate diversity since some team members would share characteristics that could be the basis for subgroup identity. Specifically creating conditions that vary the level of in-group creation would be valuable to see if the impact of virtual communication changes with the level of in-group creation.

Diversity of national background and national culture is multi-faceted and is a challenge to assess adequately. We created teams that were diverse on four aspects – country of birth, nationality, native language, and the individualism/collectivism dimension of Hofstede's measure of national culture. Although we feel that these provided reasonable indications of team-level cultural diversity, there are other aspects of national culture that could be examined (e.g., the other four Hofstede dimensions) and they could be examined singularly or together to see if complex interactions are present. For example, power distance could be important to the way team member interact, if there is diversity of status present in the team. Ethnicity could also be examined since it may be a more direct

assessment of surface-level diversity, which is the basis for the creation of subgroups early in the life of a team. A limitation of our work is that we did not directly measure ethnicity; however, people with the same birth country and language are more likely to be of the same ethnic background (although not always since many countries have multiple ethnic groups). Research determining which aspects of cultural diversity are the most important to team functioning would also be valuable. If one or two aspects were identified as being critical, companies could use this information to create a diagnostic tool used to identify good candidates for working in virtual teams and/or to identify training needs.

We deliberately studied a team that had a short life (about one hour) in order to examine if the potential for reductive capabilities of electronic media were of benefit in the early life of a team, as proposed by Carte and Chidambaram (2004). While we feel the time frame we chose was likely short enough such that a sense of team identity did not develop, future research should conduct longitudinal studies, measuring team identity, team processes and performance at multiple time points. Also, adding in the additive capabilities of additional media at different times would help identify the optimum time to do so. This would also allow a more complete testing of Carte and Chidambaram's theory since our study only deals with early team life.

The type of task could be varied as well. Future research could examine a task where cultural diversity is relevant to team performance. This would allow the net effect to be determined and see when the positive benefits of deep-level diversity overcome the negative impacts of surface-level diversity, and if these changes are consistent for teams that use different communication modes.

Our study examined one bundle of electronic communication capabilities for a team communicating synchronously and compared this to teams communicating F2F. The technology we chose was based on what we have seen in virtual team practice; however, there are technologies that have stronger reductive capabilities. Research that varies the reductive strength would be valuable since this would help identify the optimum set of technologies for use early in a team's life. For example, even though common in practice, it could be that it is better to introduce telephone conferencing at a later stage in a team's development. There are also research opportunities to vary different other aspects of virtuality (time and space) to examine if the findings are consistent for different types of virtual teams.

## Conclusion

Diversity in national background and culture is common in transnational and virtual teams. Understanding the potential advantages and disadvantages of this diversity is important for organizations. Our study focused on the potential negative aspects of diversity in national background and culture early in the life of a team and found that negative impacts appeared to be reduced by using appropriate communication media. This finding supports Carte and Chidambaram's (2004) theory that reductive capabilities in collaborative technologies can reduce the salience of surface-level diversity. Doing this early in the life of a diverse team is beneficial since it allows a team identity to form and reduces the tendency of diverse teams

to break into subgroups. Much more work needs to be done to understand diversity in teams and how to manage it effectively.

## Acknowledgements

Funding from the Social Science and Humanities Research Council of Canada and the Research Program of the School of Business at Queen's University is gratefully acknowledged. This study benefited significantly from the research assistance of Sayyida Jaffer and advice of Jane Webster. An earlier version of the paper was presented at the Academy of Management 2004 Annual Meeting held in New Orleans. We thank the audience for their attention and suggestions. We also thank the editor and three anonymous reviewers of the special issue of GDN for their suggestions which led to a significantly improved paper.

## References

Anderson, W.N. and S.R. Hiltz. (2001). "Culturally Heterogeneous vs. Culturally Homogeneous Groups in Distributed Group Support Systems: Effects on Group Process and Consensus," *Proceedings of the 34th Hawaii International Conference on System Sciences*.

Baltes, B. B., M.W., Dickson, M.P., Sherman, C.C., Bauer, and J.S. LaGanke. (2002). "Computer-Mediated Communication and Group Decision Making: A Meta-Analysis," *Organizational Behavior and Human Decision Processes* 87(1), 156–179.

Blau, P.M. (1977). *Inequality and Heterogeneity : A Primitive Theory of Social Structure*, New York, Free Press.

Bond, M.H. (1988). "Finding Universal Dimensions of Variation in Multicultural Studies of Values: The Rokeach and Chinese Value Surveys," *Journal of Personality and Social Psychology* 55(6), 1009–1015.

Bottger, P. C. and P.W. Yetton. (1987). "Improving Group Performance by Training in Individual Problem Solving," *Journal of Applied Psychology* 72, 651–657.

Carte, T. and L. Chidambaram. (2004). "A Capabilities-Based Theory of Technology Deployment in Diverse Teams: Leapfrogging the Pitfalls of Diversity and Leveraging its Potential With Collaborative Technology," *Journal of the Association for Information Systems* 5(11–12), 448–471.

Earley, C. (1989). "Social Loafing and Collectivism: A Comparison of the United States and the People's Republic of China," *Administrative Science Quarterly* 34, 565–581.

Earley, P. C. and E. Mosakowski. (2000). "Creating Hybrid Team Cultures: An Empirical Test of Transnational Team Functioning," *Academy of Management Journal* 43(1), 26–49.

Evaristo, R. (2003). "The Management of Distributed Projects Across Cultures," *Journal of Global Information Management* 11(4), 58–70.

Green, S. G. and T.D. Taber. (1980). "The Effects of Three Social Decision Schemes on Decision Group Process," *Organizational Behavior and Human Performance* 25, 97–106.

Hambrick, D. C., S. Canney Davison, S.A. Snell, and C.C. Snow. (1998). "When Groups Consist of Multiple Nationalities: Towards a New Understanding of the Implications," *Organization Studies* 19(2), 181–205.

Haslam, S. A., C. McGarty, P.M. Brown, R.A. Eggins, B.E. Morrison, and K.J. Reynolds. (1998). "Inspecting the Emperor's Clothes: Evidence that Random Selection of Leaders can Enhance Group Performance," *Group Dynamics* 2(3), 168–184.

Hofstede, G. (1980). *Culture's Consequences: International Differences in Work-Related Values*. Beverly Hills, Sage Publications.

Hofstede, G. (1983). "The Cultural Relativity of Organizational Practices and Theories," *Journal of International Business Studies* 14, 75–89.

Hofstede, G. (1994). VSM94: *Value Survey Module 1994 Manual*, IRIC, Tilberg. Accessible at http://feweb.uvt.nl /center/hofstede/VSM.html (August 2005).

James, L. R., R.G. Demaree, and G. Wolf. (1984). "Estimating Within-Group Interrater Reliability with and Without Response Bias," *Journal of Applied Psychology* 69, 85–98.

Janis, I. L. *Victims of Groupthink* Boston, MA: Houghton Mifflin, 1982.

Jehn, K. A. G.B. Northcraft, and M.A. Neale. (1999). "Why Differences Make a Difference: A Field Study of Diversity, Conflict, and Performance in Workgroups," *Administrative Science Quarterly* 44, 741–763.

Johansson, C., Y. Dittrich, and A. Juustila. (1999). "Software Engineering Across Boundaries: Student Project in Distributed Collaboration," *IEEE Transactions on Professional Communication* 42, 286–296.

Johnson, D. W. and F.P. Johnson. (1994). *Joining together: Group Theory and Group Skills* (5th ed.), Boston, Allyn & Bacon.

Kayworth, T. and D. Leidner. (2000). "The Global Virtual Manager: A Prescription for Success," *European Management Journal* 18, 183–194.

Kessapidou, S. and N.C. Varsakelis. (2002). "The Impact of National Culture on International Business Performance: The Case of Foreign Firms in Greece," *European Business Review* 14(4), 268–275.

Lau, D. C. and J.K. Murninghan. (1998). "Demographic Diversity and Faultlines: The Compositional Dynamics of Organizational Groups," *Academy of Management Review* 23(2), 325–340.

Maznevski, M. and K. Chudoba. (2001). "Bridging Space Over Time: Global Virtual Team Dynamics and Effectiveness," *Organization Science* 11, 473–492.

McDonough, E. F., K.B. Kahn, and A. Griffin. (1999). "Managing Communication in Global Product Development Teams," *IEEE Transactions on Engineering Management* 46(4), 375–386.

McGrath, J.E. (1984). *Groups: Interaction and Performance*, Englewood Cliffs, N.J., Prentice-Hall.

McLeod, P. L., S.A. Lobel, and T.H. Cox. (1996). "Ethnic Diversity and Creativity in Small Groups," *Small Group Research* 27(2), 248–264.

Milliken, F. J. and L.L. Martins. (1996). "Searching for Common Threads: Understanding the Multiple Effects of Diversity in Organizational Groups," *Academy of Management Review* 21(2), 402–433.

Mortensen, M. and P. Hinds. (2001). "Conflict and Shared Identity in Geographically Distributed Teams," *International Journal of Conflict Management* 12(3), 212–238.

Potter, R. E. and P.A. Balthazard. (2002). "Virtual Team Interaction Styles: Assessment and Effects," *International Journal of Human-Computer Studies* 56, 423–443.

Powell, A., G. Piccoli, and B. Ives. (2004). "Virtual Teams: A Review of Current Literature and Directions for Future Research," *Data Base for Advances in Information Systems* 35(1), 6–36.

Rogelberg, S. G. and M.S. O'Connor. (1998). "Extending the Stepladder Technique: An Examination of Self-Paced Stepladder Groups," *Group Dynamics: Theory, Research and Practice* 2, 82–91.

Salk, J. E. and M.Y. Brannen. (2000). "National Culture, Networks, and Individual Influence in a Multinational Management Team," *Academy of Management Journal* 43(2), 191–202.

Sarker, S. and S. Sahay. (2002). "Information Systems Development by US-Norwegian Virtual Teams: Implications of Time and Space," *Proceedings of the Thirty-Fifth Annual Hawaii International Conference on System Sciences*.

Sondergaard, M. (1994). "Research Note: Hofstede's Consequences: A Study of Reviews, Citations and Replications," *Organizational Studies* 5(3), 447–456.

Staples, D. S., I.K. Wong, and A.F. Cameron. (2004). "Best Practices for Virtual Team Effectiveness," in: David Pauleen (ed.), *Virtual Teams: Projects, Protocols and Processes*, Idea Group Publishing, Hershey PA, pp. 160–185.

Stedham, Y.E. and J.H. Yamamura. (2004). "Measuring National Culture: Does Gender Matter?" *Women in Management Review* 19(5/6), 233–243.

Straus, S.G. (1996). "Getting a clue: The Effects of Communication Media and Information Distribution on Participation and Performance in Computer-Mediated and Face-to-Face Groups," *Small Group Research* 27, 115–142.

Thompson, L.F. and M.D. Coovert. (2003). "Teamwork online: The Effects of Computer Conferencing on Perceived Confusion, Satisfaction, and Postdiscussion Accuracy," *Group Dynamics: Theory, Research and Practice* 7(2), 135–151.

Van de Ven, A.H. and D.L. Ferry. (1980). *Measuring and Assessing Organizations*. New York: John Wiley & Sons.

Van Ryssen, S. and S.H. Godar. (2000). "Going International Without Going International: Multinational Virtual Teams," *Journal of International Management* 6, 49–60.

Watson, W.E., K. Kumar, and Michaelsen. (1993). "Cultural Diversity's Impact on Interaction Process and Performance: Comparing Homogeneous and Diverse Task Groups," *Academy of Management Journal* 36(3), 590–602.

Williams, K.Y. and C.A. O'Reilly. (1998). "Demography and Diversity in Organizations: A Review of 40 Years of Research," in B. M. Staw and L.L. Cummings (eds.), *Research in Organizational Behavior* (20), 77–140.