Evaluating Self and Others in Electronic and Face-to-Face Groups

Suzanne Weisband
University of Arizona

Leanne Atwater
Arizona State University West

After completing a decision task electronically or face to face, 105 students rated their own and other group members' contribution to the task completion and their degree of liking for group members. Actual contributions were the number of task relevant remarks each person contributed. Results indicated that self-ratings of contribution were more inflated and less accurate in electronic communication than in face-to-face communication. Liking accounted for significant variance in ratings of others' contributions in face-to-face groups, whereas actual contribution accounted for significant variance in ratings of others in electronic groups. Results suggest that rating biases stemming from liking are evident in ratings of others in face-to-face groups but not in electronic. Implications for online performance evaluations are discussed.

The rise of electronic communication systems in the workplace for managing personnel, monitoring performance, conducting group work over distances, and telecommuting raises questions about how individuals perceive and evaluate themselves and others when interactions take place over computer networks. With the growing number of distributed teams communicating over computer networks, it is easy to imagine scenarios where group members and their managers interact without prior face-to-face contact, where decisions about performance or training are made without supervisors ever having observed actual job performance. Examples include software development teams (Kraut & Streeter, 1995), distance learning classes (Storck & Sproull, 1995), and other electronic groups and task forces (e.g., Orlikowski & Yates, 1994; Straus, Weisband, & Wilson, 1998) who may only meet after considerable work has been done electronically, if they ever meet at all. The implication is that information transmitted through various kinds of electronic messaging systems is being used to collect and exchange performance information. Yet, there is little systematic research about how different media affect self and other evaluations (for exceptions, see Hinds, 1997; Kulik & Ambrose, 1993; Straus, 1997).

The purpose of this study is to better understand how performance evaluations of self and others that are based on electronic interaction compare with evaluations that are based on face-to-face interaction. In particular, we examine the dimensions that raters use to evaluate themselves and others and assess whether raters' evaluations are more or less biased when interactions take place electronically, as compared with face-to-face interactions.

Background

A number of studies have investigated the ways in which electronic communication and face-to-face communication differ (e.g., McGrath & Hollingshead, 1994; Sproull & Kiesler, 1991). Some researchers claim that face-to-face communication provides richer information than electronic communication in part because computer mediation generally lacks social context cues, or social information about the members of the group (e.g., Daft & Lengel, 1986). Richer media permit people to modify their messages in response to cues from their teammates, provide immediate feedback, and allow multiple nonverbal and paralinguistic cues to confirm that people understand one another (Clark & Brennan, 1991; Daft & Lengel, 1986; Galegher & Kraut, 1994). The implication is that the additional social and nonverbal cues found in richer media would increase the extent to which messages can be accurately communicated.
and understood. However, when raters are distracted by too many social and nonverbal cues, it may provide extraneous information that interferes with the content of the message. As such, richer media may decrease accuracy and bias ratings of self and others. The present study investigates whether the lack of nonverbal behavior and reduced social cues in electronic communication effects the way people evaluate themselves and others.

**Media Effects on Self-Ratings**

Extensive research on face-to-face groups demonstrated that self-ratings are inflated when compared with ratings provided by others or to objective criteria (e.g., Atwater & Yammarino, 1997; Harris & Schaubroeck, 1988; Mabe & West, 1982). A number of factors contribute to inaccurate (inflated) self-assessments, including (a) lack of self-awareness; (b) ego defense, that is, people prefer to see themselves positively and therefore seek and attend to positive information (Miller, 1976; Swann & Read, 1981); and (c) observers tend to withhold negative feedback from others (Larson, 1986). Because nonverbal behaviors, such as gestures, head nods, facial expressions, and tone of voice, are reduced in electronic communication, the feedback individuals receive about their own behavior is limited. This reduced feedback also contributes to lowered self-awareness. With fewer social context cues, people communicating electronically feel a greater sense of anonymity and detect less individuality in others. Taken together, electronic communication lowers an individual's sensitivity to others, as well as to self, making it more likely that individuals interacting electronically will inflate their own self-ratings relative to those communicating face to face. Thus,

Hypothesis 1: Group members communicating electronically will rate their own contributions higher than will face-to-face group members.

**Likability and Performance in Ratings of Others**

Extensive research found that people who interact face to face with one another are more likely to be positively disposed toward each other (e.g., Byrne, 1971; Zander & Havelin, 1960). Even groups of strangers prefer to work with one another once they interact face to face (Zander & Havelin, 1960). As people work with others, they can observe many work behaviors and personalities. The way people form positive impressions of others is to some degree dependent on the visual and vocal behavior of the targets (Clark & Rutter, 1985). Face-to-face interaction, as a richer communication medium, offers more opportunities for revealing individuating information to group members (e.g., Weisband, Schneider, & Connolly, 1995). That is, when group members can see each other, they are likely to evaluate each other on more personal information (e.g., appearance, eye contact, head nods) than in situations where such visual cues are missing (for a review, see Fiske & Neuberg, 1990).

The importance of personal information in forming impressions of others has been demonstrated in a number of different studies (Cardy & Dobbins, 1986; DeNisi & Williams, 1988; Srull & Wyer, 1989). For example, positive first impressions (e.g., likability) about a person often lead to more positive evaluations of that person (Robbins & DeNisi, 1994). Chaiken and Eagly (1983) found that the likable communicator was more salient and persuasive when his message was transmitted by videotape or audiotape than when it was transmitted in written form. They argued that the nonverbal cues presented by the communicator in the videotape and audiotape drew participants' attention to the characteristics of the communicator (and away from the content of the message), thus making him more persuasive (Chaiken & Eagly, 1983).

The reduced social cues and awareness of others in electronic communication may help explain why group interactions appear less friendly and more serious (Culnan & Markus, 1987), especially in short-duration interactions where the time to develop interpersonal relationships is limited (Walther, 1995). Research has demonstrated that the less group members know about others, the more likely they will be to evaluate them negatively (e.g., Wilder, 1981). For example, some studies found that electronic group members reported less attraction to one another than did members of face-to-face groups (Kiesler, Zubrow, Moses, & Geller, 1985; Straus, 1997). Given the greater salience of personal information in face-to-face interactions, we expect that

Hypothesis 2: Group members interacting face to face will like each other more than will electronic group members.

Hypothesis 3: The degree to which group members like each other will account for more variance in ratings of others in face-to-face groups, compared with electronic groups.

A number of studies have shown that liking can bias performance appraisal ratings in favor of the liked person (Judge & Ferris, 1993; Katz & Glass, 1979; Tsui & Barry, 1986; Wayne & Ferris, 1990). Recent research suggests that electronic communication may reduce stereotypes and biases by filtering out irrelevant individual differences, for example, race, gender, appearance, and age (Culnan & Markus, 1987). Straus (1998), for instance, found that interviewers evaluated applicants more accurately in telephone interviews than by videoconference or face-to-face interviews. She argued that because the telephone eliminated the visual cues, interviewers were better able to focus on the applicants' responses to questions and were less concerned with or distracted by irrelevant cues (e.g., poor eye contact, fidgeting).

Consistent with Hypothesis 3 above, there is also grow-
ing evidence that rich media may lead to less accurate social judgments of others (Hinds, 1997). For example, Hinds (1997) suggested that the additional social information conveyed over a richer medium (audio-video vs. audio only) may increase the possibility of cognitive overload (see also Storck, 1995, and Straus, 1998, for similar arguments). Thus, when raters are distracted by too many social cues (e.g., maintaining eye contact, observing appropriate physical and nonverbal behaviors), cognitive load increases, and evaluations of others' contributions to a discussion are likely to be biased (e.g., Kruglanski & Freund, 1983). The result is that because more irrelevant information is available, it may be more difficult to evaluate others' contributions accurately in a rich medium (e.g., face to face) than it would be in a less rich medium (e.g., telephone or electronic mail).

Thus, by filtering out the social and nonverbal cues found in face-to-face interaction, computer-mediated communication may allow people to focus more on relevant performance information and less on irrelevant social and nonverbal information. In other words, fewer social cues may permit raters to attend to the content of the message rather than the personal characteristics of the person, leading to more accurate evaluations of group members when they interact electronically than when they do so face to face. As a result, we expect that

Hypothesis 4: Actual contributions will account for more variance in ratings of others' contributions in electronic groups as compared to face-to-face groups.

Method

Participants

One hundred five business students (64 men and 41 women) participated in a larger study examining the effects of computer-mediated communication on group influence. Participants were randomly assigned to three 3-person groups using two different communication modalities and discussing three different tasks. Each participant did one task face to face and two tasks electronically with three different sets of group members. A within-subjects factorial design was used in which communication modality and decision task were repeated. Order of communication condition was counterbalanced (i.e., some participants participated in face-to-face groups first, some participated in electronic groups first), as was the order of presentation of decision tasks. Computer and technical difficulties caused the loss of some data, but usable data were obtained from 91 groups. Of those, 27 groups met face to face, and 64 groups discussed problems over computer networks. Gender composition of the groups was not controlled, but random assignments within experimental sessions produced representative numbers of mixed-gender groups.

Decision Task

The three decision tasks used in this study required group members to make an ethical evaluation of the conduct of a computer professional in a hypothetical situation (Weis, 1990). The first situation involved the marketing of software known to have bugs, the second concerned the development and sale of marketing profiles from public information, and the third situation revolved around a university student who offered limited access to an online pornographic questionnaire. Participants rated the degree to which the central person's actions were ethical or unethical on a scale from 1 (very unethical) to 10 (very ethical).

Procedures

Participants in each session were randomly assigned to various 3-person groups prior to their arrival. All participants were first given paper copies of the experimental tasks and asked to rate privately the ethicality of the central actors in the three scenarios. On completion, the private opinions were collected, and participants then met as a group to discuss and come to consensus about the problem they were asked to do privately. If there were no questions, groups who were to first meet face to face were then asked to follow experimental assistants to their assigned rooms. Group discussions were tape-recorded, and the tapes were later transcribed for use in content analyses.

For the computer condition, each participant was assigned to one of 24 terminals in a networked electronic meeting room. The terminals were networked into eight groups of three, and group members were seated in a way that prevented them from readily determining who the members of their particular group were. To avoid verbal interaction, we made sure that participants in the same group were not seated near each other, and talking was forbidden except to ask questions of the experimenter. However, the lab facilities did not prevent participants from seeing one another. All groups were logged on synchronously to their own three-person computer conference. Messages to be sent were typed and displayed in the bottom half of the screen. Scrolling through the upper screen allowed individuals to read all previous messages. In addition, group members could read new messages while they composed their own messages, thereby reducing the exchange of redundant information. All messages were recorded automatically and later transcribed. Experimenters were on hand in both conditions to answer questions, technical or otherwise.

After the group reached consensus on the decision task, a questionnaire was distributed to evaluate participants' self-rated expertise in the problem area, their perceptions of their own and others' contributions to the final answer, and their assessments of how the group members and the communication technique they used affected their enjoyment, effectiveness, and quality of the task they completed. They were then debriefed and thanked for their participation.

Dependent Measures

Our central research interest was in self- and other-ratings as a function of a group's communication modality. At the individual level, we included ratings of group members' perceived contribution to the decision task, liking toward others, and actual contributions.

Self-ratings and ratings of others were derived from survey responses measuring each group member's contribution to the final decision. Participants were asked to "divide 100 percentage points
among all 3 group members (including yourself). Thus, ratings of self and other contributions summed to 100%. As such, ratings of self and ratings of others in each task group were inversely related.

To measure liking, we used two Likert questions ranging from 1 (strongly disagree) to 7 (strongly agree). One question asked how much participants enjoyed working with a particular member, and the second question asked participants to rate others on their desire to work with the other member again. These measures of liking have been used frequently in research on person perception and impression formation (e.g., Chaiken & Eagly, 1983). They are also similar to measures of team viability in its emphasis on desire for future contact (see Sundstrom, DeMeuse, & Futrell, 1990). The alpha coefficient for these two items when combined was .96.

Ratings of others' contributions and liking were computed as two separate measures. Because each individual (self-rater) evaluated two others on their contribution and likability, we calculated an average likability rating for the other two group members, as well as an average rating of others’ contribution to the task.

Actual contributions were derived from content analysis of the transcripts of each group’s discussion. The content coding scheme came from previous research on computer-mediated communication (e.g., Siegel, Dubrovsky, Kiesler, & McGuire, 1986; Weisband, 1992), from a modification of the Time-by-Event-by-Member Pattern Observation (TEMPO) processing coding system (Futoran, Kelly, & McGrath, 1989) and from Bales’ interaction process analysis (IPA; Bales, 1950). Actual contributions made during group discussion were measured by counting the number of task-relevant remarks an individual made during each group discussion as a fraction of the total task-relevant remarks made in that group. We defined task-relevant remarks as those classified as arguments, asking for opinions/information, agreement/disagreement, and implicit/explicit proposals. Intercoder reliability was .87 and coding differences were resolved together.

**Results**

Table 1 presents the means and intercorrelations for face-to-face and electronic groups. We first found support for Hypotheses 1: Self-ratings were significantly higher in electronic groups ($M = 0.38$) than they were in face-to-face groups ($M = 0.35$), $F(1, 257) = 4.19, p < .05$. Moreover, when self-ratings were compared with actual contributions in the two conditions, self-ratings ($M = 0.38$) were significantly higher than actual contributions ($M = 0.33$) in the electronic condition, $t(151) = 34.52, p < .001$, but self-ratings in the face-to-face condition ($M = 0.35$) were close to actual contributions ($M = 0.34$). Although evaluations of liking were high in both conditions, group members liked each other more when communicating face to face ($M = 5.71$), as compared with communicating electronically ($M = 5.03$), $F(1, 271) = 10.7, p < .01$, thus supporting Hypothesis 2.

When groups communicated face to face, how much group members liked each other was significantly correlated with their ratings of others' contributions ($r = .46$), supporting Hypothesis 3. In the electronic groups, liking was significantly related to ratings of others ($r = .15, p < .05$), but the correlation was significantly lower than that for face-to-face groups (Fisher’s $z$ test for difference between the two correlations was $2.26, p < .05$). Actual contributions were significantly correlated with ratings of others in the electronic groups ($r = .28 p < .001$), but not in the face-to-face groups, supporting Hypothesis 4. Thus, raters in face-to-face groups used liking to evaluate others, whereas raters in electronic groups relied mostly on members’ actual contributions to the group discussion to evaluate others.

When we compared the variance explained by liking in the two conditions, it accounted for 21% of the variance associated with ratings of others in face-to-face groups but only 2% of the variance in electronic groups. These findings suggest that positive affect toward others is easier to de-

---

**Table 1**

<table>
<thead>
<tr>
<th>Mode and variable</th>
<th>$N$</th>
<th>$M$</th>
<th>$SD$</th>
<th>Liking</th>
<th>SLF</th>
<th>RO</th>
<th>CS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Face to face</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liking of others</td>
<td>81</td>
<td>5.71</td>
<td>1.12</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-ratings of contribution (SLF)</td>
<td>80</td>
<td>.35</td>
<td>.08</td>
<td>- .46***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratings of others' contribution (RO)</td>
<td>80</td>
<td>.33</td>
<td>.04</td>
<td>.46***</td>
<td>-1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual contribution of self (CS)</td>
<td>74</td>
<td>.34</td>
<td>.12</td>
<td>- .13</td>
<td>.14</td>
<td>-14</td>
<td></td>
</tr>
<tr>
<td>Actual contribution of others (CO)</td>
<td>74</td>
<td>.33</td>
<td>.06</td>
<td>.13</td>
<td>-14</td>
<td>.14</td>
<td>-1.00</td>
</tr>
<tr>
<td><strong>Electronic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liking of others</td>
<td>179</td>
<td>5.03</td>
<td>1.71</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-ratings of contribution (SLF)</td>
<td>179</td>
<td>.38</td>
<td>.13</td>
<td>- .15*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ratings of others' contribution (RO)</td>
<td>179</td>
<td>.31</td>
<td>.04</td>
<td>.15*</td>
<td>-1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actual contribution of self (CS)</td>
<td>153</td>
<td>.33</td>
<td>.09</td>
<td>- .06</td>
<td>.28***</td>
<td>-28***</td>
<td></td>
</tr>
<tr>
<td>Actual contribution of others (CO)</td>
<td>153</td>
<td>.33</td>
<td>.05</td>
<td>.06</td>
<td>-28***</td>
<td>.28***</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*Note. Because self- and others' contributions (both actual and rated contributions) had to sum to 100%, self-ratings were always the inverse of others' ratings. For clarity, numbers in italics indicate redundant correlations.*

* $p < .05$. ** $p < .01$. *** $p < .001$. 

---

**RESEARCH REPORTS**

635
develop and quite influential to judgments in face-to-face discussions. In contrast, performance was significantly related to ratings of others in the electronic groups, but it only accounted for 8% of the variance. Although liking contributes very little to ratings in electronic groups, the implication is that electronic groups are rating others on much more than performance.

To gain an additional perspective, we performed regression analyses to determine the extent to which liking and actual contribution, when considered together, predicted others’ ratings in face-to-face and electronic conditions. These results are reported in Table 2. The total variance accounted for was greater in the face-to-face condition. Similar to the correlational analyses, in the face-to-face condition, liking accounted for the largest portion of unique variance, and actual contribution did not account for a significant portion of unique variance in ratings of others’ contributions. In the electronic condition, actual contributions accounted for a significant portion of unique variance, and liking did not. Total variance accounted for by liking and actual contribution was greater in the face-to-face condition than in the electronic condition.

Because there has been some controversy in the literature regarding the use of subgroup analyses, we also tested the effects of liking and contribution on ratings as a function of condition with moderated multiple-regression analyses. Media condition (Step 1), liking and actual contribution (Step 2), and the interaction of condition and liking and the interaction of condition and contribution (Step 3) were entered to predict ratings of others. (The three-way interaction was eliminated because of severe multicolinearity). The results (presented in Table 3) revealed that both interaction terms and the change in each $R^2$ were significant when the interaction terms were entered. These analyses also support our conclusion that type of communication media moderates the relationship between liking and ratings of others as well as between actual contribution and ratings of others.

**Discussion**

The purpose of this study was to investigate how face-to-face and electronic communication modalities affect self- and other-ratings. We found that (a) group members liked each other more (i.e., they were more positive toward working with one another) when communicating face to face than electronically, (b) self-ratings were higher and more inflated in electronic interactions than in face-to-face interactions, (c) liking accounted for significant variance in ratings of others in face-to-face groups but not in electronic groups, (d) actual contributions accounted for significant variance in ratings of others in electronic groups, but not in face-to-face groups, and (e) the total variance accounted for by liking and actual performance was higher in the face-to-face condition than in the electronic condition.

One apparent advantage to the reduced liking in electronic groups is the tendency for electronic group members to base their ratings of one another more on actual contributions than on feelings of liking. Biases associated with liking are more likely to influence ratings in face-to-face groups. However, there is prior work that suggests that liking of others is a result of the individual’s good performance rather than a factor that biases performance ratings (Varma, DeNisi, & Peters, 1996). Given that in this study actual contributions in the two types of groups were not different, yet liking was significantly higher in face-to-face groups, it seems unlikely that liking stemmed from performance but rather that liking served to positively bias ratings of others (or reduce self-ratings, thereby allowing greater allocation to ratings of others).

Table 3

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>$\beta$</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
<th>$F$</th>
<th>$df$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition (elec. vs. ftf.)</td>
<td>-.93*</td>
<td>.01</td>
<td>2.79</td>
<td>1,223</td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liking of others</td>
<td>-.18</td>
<td>.10</td>
<td>.09</td>
<td>10.29***</td>
<td>3,221</td>
</tr>
<tr>
<td>Actual contribution</td>
<td>-.60**</td>
<td>.10</td>
<td>.09</td>
<td>10.29***</td>
<td>3,221</td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liking X Condition</td>
<td>.78*</td>
<td>.13</td>
<td>.03</td>
<td>3.61*</td>
<td>5,219</td>
</tr>
<tr>
<td>Contribution X Condition</td>
<td>.59*</td>
<td>.13</td>
<td>.03</td>
<td>3.61*</td>
<td>5,219</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.13</td>
<td></td>
<td>.03</td>
<td>3.61*</td>
<td>5,219</td>
</tr>
</tbody>
</table>

Note. Standardized betas presented are those when all variables are in the equation. The dependent variable is ratings of others’ contributions. Elec. = electronically; ftf. = face to face.

The correlation between liking of others and ratings is common. However, Atwater, Roush, and Fischthal
Weisband, 1997). It is therefore an empirical question whether face-to-face interaction is required for trust among team members. The implication is that initial face-to-face interactions must be electronic and where trust and loyalty development of trust. However, recent research suggests that electronic teams can develop trust over time (Lacono & Others, 2000).

ratings in electronic groups. Future research to support this claim (Straus, 1998). When raters interacted face to face, impressions of others were more positive but also more biased. Managers may want to consider the purpose of performance evaluations in work settings. If the goal is to obtain accurate appraisals of others’ performance, eliminating face-to-face contact may improve accuracy. However, if it is important to obtain impressions of others that go beyond actual performance outcomes (e.g., whether the person is trustworthy, honest, funny), then face-to-face contact may be required.

There are a number of implications of this study. First, the findings suggest that in contexts where biases that are due to liking may occur and may interfere with valid judgments, electronic communication may be advantageous. For example, selection interviews are notorious for their first-impression and attractiveness biases (e.g., Cash & Janda, 1984). In some jobs for which face-to-face impressions are unimportant (e.g., telemarketing), selection interviews conducted electronically (or by telephone) could help interviewers make less biased selection decisions. For example, Straus (1998) found that telephone interviews were more accurate than those conducted face to face. Thus, if the goal is to obtain accurate appraisals of others’ performance, eliminating face-to-face contact may help.

Second, self-rating inflation appears to be more likely when group members interact electronically than face to face, suggesting that the feedback cues individuals receive in face-to-face communication may help them form more accurate self-perceptions. Or, alternatively, the liking that develops among group members reduces the individual’s tendencies to inflate self-ratings. Given the nature of the performance measures used in this study, wherein raters had to allocate contributions among self and others to total 100%, it is uncertain whether self-ratings were lowered to allot more to others or were lowered because self-perceptions were, in fact, more accurate. Future research should attempt to determine the causes of inflated self-ratings in electronic groups.

Third, there are a number of contexts where team members must interact electronically and where trust and loyalty among members is critical (e.g., software development teams, the military cohort of the future). If team members do not have the opportunity to interact face to face, it may impede the development of liking and perhaps trust among team members. The implication is that initial face-to-face meetings, even for a short duration, can enhance group members’ liking for one another and could contribute to the development of trust. However, recent research suggests that electronic teams can develop trust over time (Iacono & Weisband, 1997). It is therefore an empirical question whether face-to-face interaction is required for trust development.

Given the changes in the structure of organizations from hierarchies to flatter structures, networks, and geographically dispersed teams (DeSanctis & Poole, 1977), as well as the dramatic increases in electronic interaction among work group members, remote workers and their supervisors will not be expected to meet face to face. The result is that performance evaluations will have to be conducted without the benefit of observing behavior. Managers will then need to devise new measurement and control systems for their remote workers. Some alternative forms of performance evaluations in these virtual environments include shifting the responsibility for monitoring work progress from the supervisor to the remote employee, as well as using a team-based, peer-evaluated control system. Here, performance is evaluated according to project completion, performance outcomes, customer satisfaction, and even financial issues (Davenport & Pearson, 1998).

However, regardless of the changes in the control, measurement, and evaluation of remote workers, the organizational implications of this distributed form of work are not well understood. The results of this study suggest that distributed work arrangements affect the impressions that group members form of one another. By removing face-to-face contact, rating accuracy was improved, and there is recent research to support this claim (Straus, 1998). When raters interacted face to face, impressions of others were more positive but also more biased. Managers may want to consider the purpose of performance evaluations in work settings. If the goal is to obtain accurate appraisals of others’ performance, eliminating face-to-face contact may improve accuracy. However, if it is important to obtain impressions of others that go beyond actual performance outcomes (e.g., whether the person is trustworthy, honest, funny), then face-to-face contact may be required.

Limitations

One obvious limitation to this study is that it was conducted in an experimental setting, making it difficult to generalize to real work environments. A second limitation is that the relationships among group members were of short duration. It is possible that over time, the differences in the types of predictors of ratings of others between the two types of media may diminish (e.g., Walther, 1994). Third, the task context in which raters interacted electronically for a short time and with no face-to-face interaction may not generalize to a wide range of real-life work contexts. However, as the nature of work changes, and virtual teams become more widespread, this type of rating context will likely become more common. Fourth, ratings and measures of actual contribution were ipsative; self-ratings and ratings of others were not independent. However, in no analysis were the nonindependent measures of ratings or contribution included together. In fact, requiring raters to allocate
relative contributions may have added additional validity to the rating process because raters could not inflate ratings for all ratees. When self-ratings went up, others’ ratings had to go down accordingly.

Directions for Future Research

This study has added to our understanding of the nature of electronic and face-to-face interactions. However, much remains to be learned. For example, what factors other than liking and actual performance are contributing to ratings? Clearly only a fraction of the variance in ratings of others was accounted for by these two variables in the electronic groups. Walther (1994) has suggested that both the duration of contact and expectations for future contact influence communication patterns in electronic groups. How might factors such as the critical nature of the task or decision or sharing of personal information prior to beginning the task influence liking and future ratings of group members? Will electronic communication improve the accuracy of ratings when group members are from different disciplines and functional backgrounds? How do we assess the validity of performance evaluations in distributed work arrangements? More importantly, is it valid and ethical to evaluate others on the basis of their personality characteristics with little regard for their performance outcomes? With the increasingly pervasive uses of electronic communication in work, employee selection, and personal interactions, little is known about how impressions are formed of other group members in these distributed settings. This study is one empirical contribution that underscores the differences between how people evaluate themselves and others after communicating face to face and electronically. Future research in this area should continue to improve our understanding of the issues, as well as address the changes organizations are facing as more and more employees conduct their work away from the office.

References


Received March 9, 1998
Revision received September 23, 1998
Accepted September 28, 1998