TEAM EMOTION RECOGNITION ACCURACY AND TEAM PERFORMANCE

Hillary Anger Elfenbein, Jeffrey T. Polzer and Nalini Ambady

ABSTRACT

Teams’ emotional skills can be more than the sum of their individual parts. Although theory emphasizes emotion as an interpersonal adaptation, emotion recognition skill has long been conceptualized as an individual-level intelligence. We introduce the construct of team emotion recognition accuracy (TERA) – the ability of members to recognize teammates’ emotions – and present preliminary evidence for its predictive validity. In a field study of public service interns working full-time in randomly assigned teams, taken together positive and negative TERA measured at the time of team formation accounted for 28.1% of the variance in team performance ratings nearly a year later.

CAN TEAMS HAVE EMOTIONAL SKILLS? THE CASE OF RECOGNIZING OTHERS’ EMOTIONS

Emotion plays a central role in organizational life (Ashforth & Humphrey, 1995; Ashkanasy, 2003; Barsade & Gibson, 1998; Putnam & Mumby, 1993).
Everyday activities in the workplace are usually accompanied by feelings that serve to shape and lubricate social interactions (Fineman, 1993, 1996). Psychological theories emphasize that a key role for emotion is a tool for communication. Theorists within the social functionalist perspective have even argued that emotions evolved in part to provide an adaptive mechanism for individuals to coordinate their relationships and interactions with others (DePaulo & Friedman, 1998; Ekman, 1992; Keltner & Haidt, 1999; McArthur & Baron, 1983). We can exchange information in the form of signals about the internal states of others (Keltner & Gross, 1999), and perceiving these cues can allow people to adjust their behavior in ways that enhance the quality of interactions. The particular information exchanged through emotional cues is most often subtle and implicit, given that we use emotional cues particularly to communicate those messages that are more difficult or uncomfortable to express in a more explicit manner. It is rare to speak openly about one’s internal state, in part because development of the ability to communicate through emotional cues largely precedes higher-order explicit communication processes (Buck, 1984), and because emotional cues are sent frequently without awareness (Buck, 1984; Gross, 2001). Given the ubiquity of communicating via subtle emotional cues, this study attempts to extend previous work in this area by examining emotion communication within teams.

Emotional Expression as Communication

Communicating via emotional signals can be particularly valuable for coordinating interdependent activities and teamwork. It can provide usable information regarding the reactions, intentions, preferences, and likely future behaviors of others in the workplace (Ashkanasy, Hartel, & Zerbe, 2000; Rafaeli & Sutton, 1989; Riggio, 2001). Colleagues, customers, and other stakeholders continually attribute meaning to the expressive displays of those around them, whether they interpret correctly or misunderstand. In addition to the flow of information that is directly workplace-relevant, communication on an emotional level can also serve as an act of sharing that personalizes workplace relationships, builds mutual understanding and empathy, and provides feedback and interpersonal influence (Ashkanasy et al., 2000; Putnam & Mumby, 1993; Rafaeli & Sutton, 1989). Judgments of others’ feelings are crucial to maintaining the quality and continuation of relationships in and out of the workplace (Fineman, 1993; Izard, 2001).
Previously researchers have considered the workplace importance of emotion communication at the individual rather than interpersonal level. Emotion recognition accuracy (ERA) has been the focus of research attention for decades, with renewed attention under the umbrella of emotional intelligence (Ciarrochi, Chan, & Caputi, 2000; Davies, Stankov, & Roberts, 1998; Mayer, Salovey, Caruso, & Sitarenios, 2001; Roberts, Zeidner, & Matthews, 2001). Consistent with the value of emotional signals in providing information about others’ internal states, individual ERA has long been associated with positive outcomes in organizational settings. Typical studies use standardized tests in which participants judge the emotional content of photographs of facial expressions, audiotapes of vocal tones, and video clips of body movement, and find a positive relationship between the accuracy of these judgments and various organization-relevant outcome variables. Results have replicated across industries and positions as diverse as business executives, foreign service officers, teachers, elementary school principals, and human service workers (Rosenthal, Hall, DiMatteo, Rogers, & Archer, 1979), therapists (Campbell, Kagan, & Krathwohl, 1971; Costanzo & Philpott, 1986; Schag, Loo, & Levin, 1978), doctors (DiMatteo, Friedman, & Taranta, 1979; Tickle-Degnen, 1998), and children in academic settings (Halberstadt & Hall, 1980; Izard, 1971; Izard et al., 2001; Nowicki & Duke, 1994). These findings suggest that being able to understand others’ emotions is of highly general value to individual effectiveness.

**Emotion Recognition as a Team Competency**

The social functionalist perspective emphasizes that the benefits of emotion in the workplace arise largely in coordinating relationships and interactions. Thus, it is valuable to extend the focus on skills such as ERA beyond the individual level to the interpersonal level as well. Organizations increasingly rely on interdependence among employees as part of daily life (Wageman, 2001). Succeeding in such environments requires effective skills for coordinating with others, and emotional abilities such as ERA are likely to be among these necessary skills (Jordan & Ashkanasy, 2006).

The present study investigates whether the recognition of colleagues’ emotions can vary across teams within the same organization shortly after they are formed, even when the teams draw membership randomly from the same population of individuals. That is, can teams differ from each other meaningfully in terms of their emotional abilities – with team ERA taking on a form that is qualitatively distinct from merely aggregating the abilities
of the individuals involved? Three areas of recent research speak to the importance of raising and empirically testing this question.

First, a recent stream of research documenting the experience of group moods (for reviews, see Barsade & Gibson, 1998; Kelly & Barsade, 2001) suggests that emotion recognition skill might exist at the team level. George (1990) found that teams could have distinct affective tone, which she defined as consistent emotional reactions by members of a work group. Likewise, Totterdell, Kellett, Teuchmann, and Briener (1998) found evidence that individual moods of both nurses and accountants were influenced over time by the collective mood of nearby colleagues. Bartel and Saavedra (2000) found that many diverse types of work groups converged in their mood, assessed both behaviorally using observers and via self-reports of work group members. Further, there was significant variation among groups in the extent of this convergence (Bartel & Saavedra, 2000; George, 1990). These studies argue that the feeling states of work group members can be meaningfully considered as a team-level phenomenon. Apart from shared experiences that can lead to feeling similar emotions, the mechanism implicated for these findings has been emotional contagion (e.g., Barsade, 2002; Hatfield, Cacioppo, & Rapson, 1994), the act of catching an emotional state from another person. Emotional contagion occurs through the primitive mimicry of another person’s emotional expressions or through a conscious understanding of another person’s emotional state, either of which can lead to sharing that emotional experience. We argue that emotion recognition can be implicated as a building block for emotional contagion, in that one must be able to recognize another person’s emotional state in order to understand it consciously. Even non-conscious mimicry requires that an expression be recognizable on an implicit level, in that completely inscrutable or unintelligible expressions can be mimicked but such mimicry would not, in turn, lead to the efference process of facial, vocal, or body feedback by which individuals experience an emotion from the act of expressing it (Hatfield, Hsee, Costello, & Weisman, 1995; Zajonc, Murphy, & Inglehart, 1989). Taken together, this suggests that teams in which members can better understand each other’s emotional states may come to share those emotional states more strongly, and that evidence for variability in the convergence of mood across groups suggests there may be variability in the mutual understanding of teammates’ emotional states.

A second area providing support for examining the emotion recognition skill of teams comes from researchers interested in team competencies, norms, and effectiveness. Flynn and Chatman (2003) have argued that group norms can develop early in the life of teams, starting from dyadic
perceptions through a process whereby teammates make judgments of each other, and these judgments trigger behavioral responses that iterate among members ultimately to form group-wide norms. In as much as group norms include a wide range of rules to guide behavior as a means to enhance group functioning and survival, to make members’ behavior more predictable, to avoid embarrassing interpersonal situations, and to express core values (Feldman, 1984), patterns of communication can be central to group norms. Correspondingly, the exchange of information more generally has been considered a source of team-level competencies (e.g., Hambrick, 1994) because, although it is often pairs of individuals and sub-groups that directly exchange information, the dyads within teams tend to reciprocate teammates’ behaviors, and mutually influence each other. Likewise, Polzer, Milton, and Swann (2002) found that the mutual ability of dyads within groups to understand each other’s social identities varied meaningfully across work groups. Druskat and Kayes (1999) argued that effective interpersonal behaviors are a form of team-level competency, and they argue in particular that these behaviors include the perception and interpersonal understanding of members’ spoken and unspoken feelings. Using interviews and questionnaires at a large manufacturing firm, they found evidence of significant variation across teams in these interpersonal skills. Just as there can be general communication and interpersonal skills that vary across teams, we argue that team-level skills can exist specifically for the exchange of information via emotional channels.

A final related area of research suggests that ERA has emergent properties beyond the individual level. Early research in this area focused exclusively on ERA as a stable trait, using standardized tests in which skill was assessed by accuracy in perceiving consistent stimuli consisting of the expressions of strangers (Rosenthal et al., 1979). However, this may not be the same as accuracy in perceiving the actual colleagues with whom one is working, because interaction partners are not simply interchangeable. There is some evidence that people may learn to identify more accurately the specific emotional expressions of relationship partners early on (Sabatelli, Buck, & Dreyer, 1982) – even though there is no evidence for further increases in accuracy within long-term relationships (Ansfield, DePaulo, & Bell, 1995; Sabatelli, Buck, & Dreyer, 1980) – which would suggest that teammates may be better at understanding each other’s emotional expressions than those of strangers. Because the communication of emotion is inherently interactive, a view of emotional skill that incorporates only stable individual differences across perceivers is incomplete. Indeed, Elfenbein, Foo, Boldry, and Tan (2006) found that – in addition to individual
differences in the ability to perceive emotional expressions, and individual differences in the complementary ability to express one’s emotions clearly to others – there can also be unique “relationship” effects as well. That is, in judging each other’s emotional expressions, some previously unacquainted dyads systematically appeared to “click,” so that the perceiver understood the expressor’s emotions better than the skill levels of the individuals would have predicted. Likewise, other dyads appeared to “cross wires,” so that the perceiver systematically failed to understand the expressor in spite of the general skill of the individuals involved. Styles of emotional expression can vary idiosyncratically – without necessarily being “better” or “worse” – and from the beginning of a relationship we may be able to judge more accurately those expressions appearing in a familiar style. These findings suggest that dyads as well as individuals can be more or less emotionally intelligent, which argues for the importance of examining emotional abilities beyond the individual level. Groups may have coherent emotional skills – including ERA – that may be more than the sum of its parts.

Team Emotion Recognition Accuracy

Our recent work established evidence that teams can have meaningful differences in a particular emotional skill, which emerge at the group level as more than a mere aggregation of skill of the individuals involved. The act of communicating does not occur alone; an audience is always implied (Goffman, 1956). This distinguishes the understanding of emotions from other emotional abilities such as emotional regulation. Thus, accuracy in understanding colleagues’ emotional states is a particularly relevant emotional skill to examine in a team context. Moving beyond the individual level examined by past researchers requires new concepts as well as empirical work to validate these concepts, which is the goal of the current study. We introduce the construct of Team Emotion Recognition Accuracy (TERA), which we define as the ability of teammates to recognize emotions expressed by their colleagues.

Based on past research, there are two potential mechanisms from which to expect that teams could vary systematically in their degree of TERA – even if the teams do not differ in the general accuracy of their members in perceiving emotions. The first mechanism is that some teams may have members who match or mismatch with each other in terms of their emotional expressive style before they even meet, and thus who may understand each other better or worse than their individual skills would predict
The second mechanism is that team members may be able to learn about each other’s emotional styles when they first meet (Sabatelli et al., 1982), and their interest and ability in doing so could vary meaningfully across work groups. Druskat and Kayes’ (1999) and Polzer et al.’s (2002) findings of meaningful variation across teams in skills of interpersonal effectiveness and understanding would imply that teams’ members could vary systematically in their degree of learning of each other’s emotional styles. Teams with healthy interpersonal norms may encourage greater mutual understanding among colleagues, whereas teams with less-effective interpersonal norms may discourage the attention and motivation to mutually understanding colleagues’ internal states. Both of these mechanisms could lead to TERA that is different from the mere aggregation of individual abilities.

It is worthwhile to distinguish this new concept of TERA from the existing concept of group mood. For example, in the course of its work, a product development team may learn about the rejection of an important proposal. Due to this shared event, as well as emotional contagion from those with strong reactions, the group mood might involve sadness and disappointment. By contrast, TERA is a meta-concept referring to whether team members are aware of how their colleagues feel – rather than the extent to which they feel the same way. In this sense, TERA shares a commonality with the concept of team transactive memory (e.g., Moreland, 1999), the sharing of mental models among colleagues (Austin, 2003). In the project team example, high TERA might mean that members know who among them are the most disappointed, versus who might even be somewhat relieved.

Requirements of Team Emotion Recognition Accuracy

In addition to the above arguments that TERA is coherent theoretically at the team level, there are certain empirical requirements that it must meet. For TERA to be a group-level construct, it must satisfy two criteria (Kenny & La Voie, 1985). The first requirement is that TERA must converge across group members. A construct cannot be said to describe the team unless members of each team are more similar to each other than they are to the members of other teams. Second, TERA must be conceptually meaningful at the group level. That is, TERA must be a characteristic of the team and not just describe the individuals on the team. In order to establish this distinction, we demonstrate empirically that TERA as a group-level
The construct is distinct from its individual-level counterpart. That is, in addition to TERA, the present study also examines General Emotion Recognition Accuracy (GERA), which is the ability to understand emotions that are expressed by non-teammates. TERA differs from GERA in that TERA focuses on accuracy at understanding teammates whereas GERA focuses on accuracy at understanding expressors in general. Both TERA and GERA can be measured at the individual level and then aggregated to the team level, with different conceptual meanings. Conceptually, team-level GERA is the equivalent of aggregating teammates’ individual-level scores on the type of standardized emotion recognition measures typically used in previous research. Conceptually, by contrast, team-level TERA is not a mere aggregation of individual scores, but rather a measure of mutual understanding within the team. Thus, TERA meets the requirement of describing the team rather than its individuals. Empirically, then, it is important to demonstrate that TERA converges among team members to a greater degree than GERA, which we do not expect to converge among team members (Fig. 1).

The goal of this study is to introduce the construct of TERA, and to document its validity with respect to these two empirical requirements. In particular, we hypothesize that there is significant convergence among teammates in TERA, which is greater than any convergence observed in GERA or other strictly individual difference concepts.

**A. Team Emotion Recognition Accuracy (TERA)**
- Judges and Targets: Teammates

**B. General Emotion Recognition Accuracy (GERA)**
- Targets: Non-Teammates

![Fig. 1. Representation of Team and General Emotion Recognition Accuracy.](image)
Team Emotion Recognition Accuracy and Team Performance

Team Outcomes and Emotional Valence

TERA should facilitate effectiveness, both in terms of performance as well as promoting a more positive social environment, because it allows more effective communication and interpersonal coordination. Making the interactions among colleagues smoother and more predictable should improve the quality of those interactions. For example, team members who understand each other’s emotional reactions may be able to refrain from behaviors that irritate or upset their colleagues, and may be able to increase the behaviors that impress or amuse them. By contrast, team members who cannot understand their colleagues’ emotions may miss the opportunity for such feedback that helps them to adjust their behavioral repertoire. Thus, the second goal of this study is to test the impact of TERA on workplace performance and social outcomes, with the hypothesis that greater TERA facilitates greater team effectiveness and cohesion among colleagues.

The social functionalist theoretical framework that grounds the above hypothesis argues for their applicability across the various emotions that individuals can experience, express, and recognize in their workplace colleagues. After all, the subtle cues relevant to coordinating interpersonal interactions can take many forms, and usable information via emotional signals can be transmitted via many different emotions. However, other streams of research on emotion in the workplace also emphasize the importance of distinguishing between positivity and negativity when studying the associations between emotional tendencies and performance.

In the case of emotional expression, in many environments positive emotion is particularly encouraged in the workplace, along with the suppression of negative emotion. For example, those who self-report more positivity in their emotional traits generally experience superior workplace outcomes (e.g., Staw & Barsade, 1993; Staw, Sutton, & Pelled, 1994; Wright & Staw, 1999). Staw et al. (1994) theorized that such positive emotions have positive consequences due to an association with greater productivity and persistence, as well as impacting employees’ relationships with colleagues. That is, feeling and expressing positive emotions on the job can lead to smoother social interactions, more helping behaviors, and a favorable “halo effect.” Further, research on emotional contagion (e.g., Hatfield et al., 1994) demonstrates that people tend to mirror the emotional states and behaviors of those around them. Taken together, these research streams suggest that greater accuracy with positive and negative expressions of emotion could lead team members to experience and express positive and negative feelings and behavior, respectively, possibly also generating positive and negative
workplace consequences, respectively. For this reason, analyses are conducted separately by emotional state, with a particular focus on the possible distinction between patterns for positive and negative emotion.

**METHOD**

A series of interviews and judgment exercises, conducted with public service interns beginning a full-time job program, assessed the accuracy of team members in understanding each other’s emotions (TERA). Participants also judged emotions expressed by non-teammates in order to measure the general emotional accuracy (GERA) of the individuals on each team.

**Participants**

Participants were employed full-time for one year by a non-profit organization that provides community services in underprivileged neighborhoods. The bulk of its work is conducted in teams, for which it assigns individuals using block randomization, blocking along gender, ethnicity, and city versus sub-urban upbringing. Members were young adults between 17 and 23 years of age, working interdependently in teams to perform a variety of functions such as serving as assistant teachers, after-school and day-camp counselors, disaster relief workers, assistants to local charities, and other public service roles working mostly with “at-risk” groups. Participants were unacquainted with teammates before the program began. Of the 114 members enrolled in the program, 68 members (60%) across 14 teams completed the majority of measures described below and were included in analyses. There were no systematic differences between these 68 members and the larger group of 114 in demographic composition (gender $t(112) = 1.218$, ns; ethnicity $\chi^2(4) = 6.997$, ns; age $t(92) = .361$, ns, for those with age data available).

**Emotions**

This study includes the four basic emotions that previous investigations of ERA have had in common – anger, fear, happiness, and sadness (e.g., Ekman, 1972; Izard, 1971; Nowicki & Duke, 1994). The study also includes embarrassment, which has been implicated for particular relevance to workplace life due to its role in maintaining social order and appeasement.
for transgressions across social convention (Fineman, 1993, 1996; Goffman, 1956; Keltner & Buswell, 1997; Tangney, 2003). Given that the experience of embarrassment can range from light-hearted humor to deep humiliation and shame, an additional pilot test examined the scenarios generated by the interview protocol outlined below. The scenarios were considered significantly more amusing than either humiliating, \( t(19) = 2.58, p < .02 \), or shameful, \( t(19) = 3.46, p < .003 \), by three research assistants rating the full-length segments from a random sample of 20 (24%) of the 82 interviews (inter-rater reliability = .77). Thus, the emotion is relabeled “Amused/Embarrassed.” In light of extensive theory concerning differences in organizational outcomes for positive versus negative emotion (e.g., George, 1990; Isen & Baron, 1991; Staw et al., 1994), results are presented below for each emotion individually, and aggregated for positive (happiness, amused/embarrassed) and negative (anger, fear, sadness) emotional categories.

Phase 1: Individual Interviews

Either during the month before the program began or within the first week of orientation, 82 employees (72%) participated in an individual videotaped interview protocol designed to provide examples of affective displays using a validated method of emotion induction. Eliciting emotion through the guided recall of past events from participants’ lives has become a common and well-tested method in research on emotional expression and experience, and has been extensively validated with both clinical and non-clinical populations as an effective elicitor of emotional responding. The guided recall of past highly emotional events reliably predicts self-ratings of emotional experience (Dunn & Schweitzer, 2005; Lerner & Keltner, 2001; Levenson, Carstensen, Friesen, & Ekman, 1991; Strack, Schwarz, & Gschniedinger, 1985; Tiedens & Linton, 2001; Tsai, Chentsova-Dutton, Freire-Bebeau, & Przmys, 2002), facial expressions and autonomic nervous system activity corresponding to the emotion recalled (Levenson et al., 1991; Tsai et al., 2002), and effects on factors such as optimism and decision-making styles that correspond to dispositional findings for the same emotions (Lerner & Keltner, 2001). Thus, guided recall appears to provide an authentic yet controlled method for emotion elicitation.

Accordingly, the initial interviews followed a protocol in which participants described and were asked to relive real experiences that had occurred in an organizational setting, either in a previous workplace or in school. To provide an opportunity to select past experiences without time pressure,
participants first completed a survey prompting them to recall experiences during which they had felt each emotion very strongly, while at work or at school, with the order of emotions varying between participants according to a balanced Latin Square. The interviewer instructed participants to look at a large sign with the name of the emotion, which positioned them to face a video camera while speaking. In order to eliminate non-verbal feedback and to limit interaction, the interviewer sat behind the interviewee, outside of the camera’s range, and was the only other person present in the room. In keeping with Strack et al.’s (1985) finding that the recall of past events triggers greater affect when recalled vividly and in detail, and when discussing how rather than why events occurred, the interviewee was asked to discuss each experience in detail, to repeat what he or she actually said and did at the time, and to use first-person language while trying to put himself or herself back into the situation.

Although there has traditionally been a distinction between the spontaneous expression of emotion versus socially conscious displays (e.g., Ekman, 1972), more recent perspectives such as Fridlund’s (1994) behavioral ecology theory of emotion have argued that expression at its core is a social phenomenon, over which we have a continuum of intentional control. Indeed, research participants in truly non-social spontaneous settings are frequently non-expressive (Fridlund, 1991, 1994). Thus, the current method was intended to elicit samples of emotional expression similar to those that participants might actually choose to display in their workplace settings.

Interviews lasted an average of 10–15 min. Participants were asked not to discuss the content of their interviews with teammates until after the subsequent phases of the study.

Phase 2: Creation of Video Stimuli

Stimuli consisted of 5-s video clips that illustrated each interviewee’s style of expressing each emotion.

Choice of Video Clips
Due to variability in the amount of time during which interviewees described their emotional events and repeated their reactions, some judgment was necessary in selecting the most representative 5-s portion. The criterion was to select the first feasible section after participants began their first-person description. Sections were unfeasible if they included revealing verbal content, which was defined as words that (1) used the name of the emotion,
(2) provided sufficient context for another person to identify what emotion they would likely feel, or (3) violated the privacy of the interviewee through sensitive language or proper names. For example, if an interviewee said, “I am so happy because of what just happened,” the video clip could include the sentence removing the single word “happy,” both the audio and video portions to prevent lip reading. Using these criteria, the first author selected the video clips from the interviews. In order to assess the reliability of this process, a research assistant also coded 25 emotional expressions from five different interviewees. Inter-rater reliability was .82, which increased to .90 upon re-coding of one segment for which one judge provided a second choice that was close to the segment chosen by the other judge.

The goal of this study is to examine the communication of emotion as it naturally occurs in organizational settings. Emotion can be expressed through many cues – ranging from strictly verbal cues to non-verbal cues such as facial expressions, vocal tones, and body language (Ekman & Friesen, 1969). Past research on the communication of naturalistic emotion has excluded explicit verbal content from stimulus materials, yet included implicit verbal content that may provide information about an expressor’s emotional state (e.g., Archer & Akert, 1977; Costanzo & Archer, 1989; Mayer & Geher, 1996). Rather than eliminate verbal content entirely through methods that disrupt the ecology of naturalistic emotional expression – such as still photographs (e.g., Izard, 1971), silent videos or filtered vocal samples (e.g., Rosenthal et al., 1979), or standardized scripted verbal content intended to provide paralinguistic cues but not information from the words themselves (e.g., Nowicki, 2000) – we sought to preserve as much as possible of the participants’ original expressions and minimize only explicit verbal cues.

A pilot test confirmed that the verbal content remaining in the video clips did not provide sufficient information alone to judge the intended emotion. Six research assistants coded written transcripts of the 5-s video clips, rating how plausible it would be for each transcript to refer to a story about each of the five emotions – afraid, angry, embarrassed, happy, and sad, using a scale from 1 to 5 (reliability = .77). Consistent with the inclusion of implicit verbal content, the judges did in fact rate the correct responses as more plausible than the incorrect responses, $t(358) = 7.30$, $p < .01$, $r = .36$. However, judges could not use the verbal information alone to select the correct response, because upon removing for each video clip the item rated as the least plausible, the correct response was no longer rated as more plausible than the remaining three choices, $t(358) = 1.58$, ns, $r = .08$. That is, given the verbal content alone, each video clip had on average four choices that
were equally plausible. Thus, participants could not achieve accurate responding based on the verbal content alone, but they could use it to rule out one implausible choice for each clip. These pilot findings are consistent with our goal to minimize only explicit but not implicit verbal cues.

Phase 3: Judgments of Video Clips by Colleagues

Team members participated in a judgment exercise one week after the teams formed, which was one week after the last interviews had taken place. At this point in the program, participants had been employed with the organization for less than two weeks of orientation, of which they knew their team assignments for one week. Although this was not long in terms of the one-year lifetime of these intact teams, the 40h of team-based training sessions and team-building activities represent a substantial amount of contact compared with teams studied in laboratory settings. Members were beginning to grow acquainted but had not yet begun their work together. A total of 82 employees (72%) participated in the judgment exercise.

Non-Teammate Segment
As a measure of general emotional accuracy (GERA), participants viewed video clips from a standardized sample of members from other teams. The non-teammate segment contained the video clips from all five emotions for 24 randomly selected participants, for an average of one to two members from each team. Block randomization, whereby each expressor appeared once before any expressor was repeated, reduced the possibility that judges could use process-of-elimination based on their memory of the emotions expressed in earlier clips. The order of emotions varied randomly. To reduce possible order effects due to practice or fatigue, the non-teammate segment was divided into two sections. Half of the teams viewed the two sections in one order and half viewed the other order. On average, in viewing this common sample, members of each team made judgments of one to two of their own teammates. However, such ratings were removed from the non-teammate data, and instead were analyzed with the teammate ratings described in the following section.

Teammate Segment
As a measure of TERA, participants viewed the video clips for each team member who had completed the emotion interview. Teammates who had already appeared in the non-teammate segment were not repeated in the
teammate segment, which was distinct for each team. Video clips were block randomized and appeared in a random order. In order to reduce possible effects of stimulus order, the teammate segment appeared in between the two sections of the non-teammate segment.

Procedure and Scoring
Participants were instructed to make their best guess of the emotion discussed in each video clip, and to circle this judgment on a multiple-choice response sheet that listed the five choices: afraid, angry, embarrassed, happy, and sad. After each video clip, participants had 5 s to enter their judgments. Participants were told that they would see each emotion for each person, so that there was an equal chance that each response would be the correct answer. Video clip responses were scored as correct if the judge selected the emotion that the interviewee had been discussing during the portion of the interview from which the video clip was selected, which indicated that the judge perceived in the display the same emotion the expressor discussed. Accuracy values were then corrected for simple response bias. Otherwise, Pollyannish participants choosing positive responses for all video clips would have high accuracy with positive emotions and low accuracy with negative emotions – and dysphoric participants the opposite. Using Wagner’s (1993) correction formula eliminates the possibility that multiple-choice accuracy values reflect mere response bias. The formula multiplies accuracy values by 1 minus the rate of false alarms, in which a category is endorsed when it did not actually appear, and then normalizes this value using an Arcsine transformation. This correction is similar to signal detection methods except, unlike signal detection terms, it allows separate analyses for each emotion (Wagner, 1993). Instances of individuals rating their own emotional expressions were excluded (Kenny & La Voie, 1984).

Phase 4: Outcome Measures

In order to assess performance and social outcomes at the end of the year-long work program, participants and staff members at the organization completed a number of questionnaires. Although there is no single meaning that is commonly accepted for effectiveness within work groups, in general researchers agree that it is a multifaceted concept incorporating multiple perspectives and metrics (Goodman, Ravlin, & Schminke, 1987; Jones, 1997).
Individual Performance Ratings
At the completion of the program, the senior staff members provided performance ratings for the participants. Fourteen staff members completed these ratings, using a scale from 0 (extremely poor) to 10 (extremely great), and they left a blank response when they did not feel sufficiently acquainted to provide a judgment for a particular individual.

Inter-rater reliability, the average product–moment correlation between raters (Rosenthal & Rosnow, 1991), which is equivalent to the ICC(1) (Bartko, 1976) was .38, and the Cronbach’s α for ratings across staff members, equivalent to the ICC(3) (Bartko, 1976) was .84. In addition to assessing individuals, staff members also rated the overall quality of the work conducted by the team. An average of 9.4 staff members rated each team. The inter-rater reliability, the average product–moment correlation between raters, was .55, and α for the composite was .92. In addition, between one and two weeks before the completion of the program, all remaining participants provided ratings for each individual who had been a part of their team over the course of the year. Sixty-three participants from the 14 teams completed these measures, for an average of 4.5 per team. Using a scale from 0 (not at all) to 10 (very much), team members rated to what extent they believed each individual had done a good job. The average Cronbach’s α across teams was .93. The correspondence of $r = .61$ between staff and teammate ratings of individuals compares favorably with previous research documenting the average correspondence among supervisor and peer performance ratings (Facteau & Craig, 2001; Furnham & Stringfield, 1998; Harris & Schaubroeck, 1988), and the staff and peer ratings combined to form a composite performance rating with a reliability of Cronbach’s α = .76.

Team-Level Outcome Variables
It was also necessary to assess the performance of each team as a whole. Hackman (1987) details a normative model of group effectiveness that focuses on the complexity faced by real teams in real organizations. Because objective quantities can rarely measure most real groups’ output, Hackman argues that organization members and clients should provide their own assessments of whether work group output meets or exceeds performance standards. Therefore, in addition to team ratings by staff provided by the same staff who rated individuals – and who had extensive contact with both the team members as well as the teams’ public service clients – participants themselves also gave self-reported ratings covering aspects of their team’s effectiveness. Between one and two weeks before the completion of their year-long job program, all remaining participants completed a short survey
that included questions based on Hackman’s theory. On a 7-point scale ranging from strongly disagree to strongly agree, team members rated: whether their team accomplished their community service goals (*Hackman Criterion 1*), and whether generally, in their opinion, the team had done a good job (*Member evaluation of team outcome*). Hackman (1987) further argues that effective teams have social processes that enhance the interest of members to work together on subsequent tasks, and that they satisfy rather than frustrate personal needs of group members. Therefore, using the same scale, team members rated whether they would want to work together again with the same team (*Hackman Criterion 2*), whether they had grown personally (*Hackman Criterion 3*) from their team experience, and their interpersonal cohesion in the form of liking for each member of their team (*liking*). Hackman’s model encompasses the major components of organizational effectiveness theory, particularly group performance, adaptiveness, and concern for multiple constituents of the team’s work (Jones, 1997). Sixty-one of the 82 participants completing the year-long program (74%) completed these ratings. After excluding two teams as described above, a total of 57 participants from the remaining 14 teams completed the members’ ratings of team outcomes, for an average of 4.1 per team.

The average inter-item reliability among these six measures of team effectiveness was $r = .26$. However, Hackman Criterion 3 had a low average inter-item correlation of $r = -.07$, and upon its removal the average inter-item reliability among the remaining five measures was $r = .42$, for a total reliability of .79 for overall team effectiveness.

*Control Variables and Concurrent Validity*

Participants completed the 60-item version of the NEO-PPI five-factor personality scale (Costa & McCrae, 1992) and provided their age, sex, and ethnicity. In order to argue that TERA is not driven merely by team differences in other attributes, we also present team convergence data on these personality traits and demographic background variables that can be related to accuracy in the communication of emotion (Elfenbein & Ambady, 2003; Hall, 1978; Izard, 1971; Nowicki & Duke, 2001; Rosenthal et al., 1979).

In order to establish that the present exercise was valid as a measure of ERA, individual general emotion accuracy (GERA) scores from the exercise were compared with those on the popular and well-validated Diagnostic Analysis of Nonverbal Accuracy (DANVA; Nowicki, 2000). The Adult Faces and Adult Paralanguage scales of the DANVA each contain 24 posed
emotional expressions using photographs of faces and audiotapes of voices, respectively. Each scale takes approximately 5 min to complete. Sixty-three of the 100 individuals on the 14 teams completed the DANVA, for an average of 4.5 members per team (63%). The average scores, 78% ($M_{18.7}$ out of 24, $SD_{3.3}$) on the Faces scale and 72% ($M_{17.3}$ out of 24, $SD_{2.4}$) on the Paralanguage scale, were similar to national norms for the instrument (Nowicki, 2000). DANVA scores showed excellent concurrent validity, with a correlation of $r = .46$ ($p < .01$) with accuracy in the non-teammate section on the video clip exercise assessing GERA. The strength of this correlation provides evidence that the current findings are robust against the potential limitations inherent in using an exercise rather than validated instrument to assess ERA.

RESULTS

Operationally, we define TERA for each individual as the average accuracy of emotion judgments of their own teammates’ expressions. By contrast, we define GERA – consistent with past research on individual differences in emotion recognition – as the average accuracy of an individual’s emotion judgments of a common set of expressors. Thus, TERA specifies a specific relationship between the perceiver and target – that of teammates – whereas GERA is defined at the individual level independently of the target, averaged across multiple standardized targets. Table 1 lists correlations among outcome, control, and ERA values.

To demonstrate convergence among teammates in their level of TERA, we used the intraclass correlation (ICC) statistic, which conceptually refers to the average correlation consisting of pairs randomly selected within the same teams (Kenny & La Voie, 1985; Kenny, Kashy, & Bolger, 1998). Table 2 presents intraclass correlations for TERA. Consistent with TEA’s status as a team construct, there is positive convergence across teammates for both positive and negative emotion. Consistent with GERA’s status as an individual-level construct, there does not appear to be positive convergence across teammates and, further, convergence is significantly greater for TERA than for GERA. Additional analyses illustrated that the current findings of team-level convergence for TERA are not likely to result from confounding team-level differences in other factors such as personality and demographic variables. There is no evidence for positive convergence along any of the big five personality traits, gender, age, or ethnic background.
Table 1. Means, Standard Deviations, and Correlations Coefficients for Outcome, Control, and Emotion Recognition Accuracy Values ($N = 68$).

| Variable                           | M    | SD   | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  | 11  | 12  | 13  | 14  | 15  | 16  | 17  | 18  | 19  |
|------------------------------------|------|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| **Outcome variable**               |      |      |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Performancea                       | .93  | .95  |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| **Control variables**              |      |      |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Female                             | .47  | .50  | .09 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| White                              | .44  | .50  | .26 | .23 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Age                                | 19.73| 2.00 | .21 | .18 | .13 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| **General emotion accuracy (GEA)** |      |      |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Amused                             | .156 | .111 | .23 | .01 | .29 | .19 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Happy                              | .160 | .084 | .10 | .07 | .14 | .05 | .04 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Afraid                             | .125 | .084 | .25 | .04 | .12 | .17 | .27 | .11 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Angry                              | .191 | .107 | .31 | .17 | .18 | .15 | .28 | .25 | .16 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Sad                                | .142 | .090 | .08 | .26 | .05 | .09 | .11 | .03 | .21 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Positive                           | .158 | .071 | .24 | .03 | .15 | .18 |     |     |     | .27 | .36 | .09 |     |     |     |     |     |     |     |     |     |     |     |
| Negative                           | .113 | .047 | .31 | .24 | .18 | .10 | .26 | .25 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Total                              | .155 | .054 | .35 | .12 | .20 | .13 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| **Team emotion accuracy (TEA)**    |      |      |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Amused                             | .269 | .265 | .04 | .09 | .00 | .22 | .08 | .11 | .19 | .12 | .26 | .12 | .27 | .26 |     |     |     |     |     |     |     |
| Happy                              | .267 | .220 | .01 | .08 | .03 | .26 | .33 | .20 | .30 | .40 | .34 | .44 | .29 |     |     |     |     |     |     |     |     |
| Afraid                             | .146 | .166 | .23 | .05 | .23 | .04 | .06 | .21 | .03 | .13 | .01 | .00 | .04 | .12 | .03 |     |     |     |     |     |     |     |
| Angry                              | .233 | .302 | .03 | .15 | .11 | .04 | .04 | .00 | .02 | .09 | .06 | .06 | .23 | .03 | .16 |     |     |     |     |     |     |     |
| Sad                                | .224 | .184 | .06 | .12 | .06 | .17 | .10 | .17 | .05 | .16 | .18 | .15 | .19 | .22 | .30 | .23 | .10 |     |     |     |     |
| Positive                           | .268 | .204 | .03 | .02 | .15 | .19 | .25 | .23 | .19 | .35 | .30 | .37 | .42 | .00 | .07 | .18 | .32 |     |     |     |     |
| Negative                           | .151 | .102 | .05 | .20 | .01 | .02 | .10 | .10 | .07 | .06 | .13 | .14 | .16 | .28 | .19 |     |     |     |     |     |     |     |
| Total                              | .226 | .121 | .04 | .13 | .05 | .11 | .20 | .22 | .17 | .29 | .29 | .33 | .38 | .30 |     |     |     |     |     |     |     |

*Note: N = 66 for liking, N = 61 for age, N = 68 for all other variables. Correlations not listed between composite measures and underlying variables contributing to the composite.

*Performance is a composite variable consisting of the average of standardized staff and peer performance ratings.

*All emotion recognition accuracy values are listed in arcsine-transformed unbiased hit rates (Wagner, 1993).

~p < .10.
*p < .05.
**p < .01.
***p < .001.
Table 2. Descriptive Statistics and Intraclass Correlations (ICC) for Judgment Accuracy in Recognizing the Emotional Expressions of Teammates and Nonteammates.

<table>
<thead>
<tr>
<th>Positive Emotion</th>
<th>M</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TERA</td>
<td>GERA</td>
</tr>
<tr>
<td>Amused/Embarrassed</td>
<td>.269</td>
<td>.156</td>
</tr>
<tr>
<td>Happy</td>
<td>.267</td>
<td>.160</td>
</tr>
<tr>
<td>Total</td>
<td>.268</td>
<td>.158</td>
</tr>
<tr>
<td>Negative Emotion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afraid</td>
<td>.146</td>
<td>.125</td>
</tr>
<tr>
<td>Angry</td>
<td>.233</td>
<td>.191</td>
</tr>
<tr>
<td>Sad</td>
<td>.224</td>
<td>.142</td>
</tr>
<tr>
<td>Total</td>
<td>.201</td>
<td>.152</td>
</tr>
<tr>
<td>Total</td>
<td>.226</td>
<td>.155</td>
</tr>
</tbody>
</table>

Note: Mean judgment accuracy listed in unbiased hit rates (arcsine transformed). TEA refers to Team Emotion Recognition Accuracy, GERA refers to General Emotion Recognition Accuracy. t refers to the statistical test of the difference between judgment accuracy for TERA and GERA. r refers to the effect size for the test of the difference between ICCs for judgment accuracy for TERA and GERA.

~p < .10.
*p < .05.
**p < .01.
***p < .001. N = 68 individuals, 14 teams.
We predicted that differences in TERA would be predictive of greater team effectiveness. First, for illustrative purposes, Table 3 lists the results of multiple regression analyses predicting workplace performance as a function of control variables and GERA. Because GERA did not show team-level effects, these analyses take place most appropriately at the individual level of analysis (Chan, 1998). Dummy variables account for membership in teams. Consistent with past research on ERA in the workplace, these analyses document a positive association between GERA and workplace accuracy, which is consistent across positive and negative emotions.

Table 4 lists correlations among TERA and team-level outcome variables. Because TERA showed significant team-level effects, these analyses most appropriately take place at the team level of analysis (Chan, 1998). In keeping with the process used for team assignment, which was random subject to maximizing diversity within teams, as described above there were no positive intraclass correlations among control variables across the 14 teams. Thus, in order to preserve statistical power, tests of the relationship between team-level workplace outcomes and the ability to communicate emotion present zero-order correlations and do not include these control variables, as they do not differ meaningfully across teams. In spite of the low statistical power at the team level, these analyses demonstrate a consistent and suggestive trend: greater TERA with positive emotion appears to predict positive outcomes, whereas greater TERA with negative emotion appears to predict negative outcomes. This trend is apparent both when examining coefficients one at a time, as well as in regression models including coefficients for both positive and negative TERA, which yield a number of significant coefficients and model diagnostics. This trend contrasts with the impact of GERA, described above, which was beneficial across types of emotion. Two more focused tests of this post-hoc comparison are also statistically significant. First, there was a significant coefficient for the positive valence effect, the extent to which the team was more accurate in expressing positive rather than negative emotion. This positive valence term was included in linear regression models also controlling for the main effect of total TERA. Second, direct tests of overlapping correlation coefficients (Meng, Rosenthal, & Rubin, 1992) confirmed significant differences between the coefficients for positive and negative emotion.

DISCUSSION

This study illustrates the potential value of considering the emotion recognition skill of teams, which is conceptually and empirically separate from the
Table 3.  Multiple Regression Models Predicting Individual Performance From General Emotion Accuracy in Recognizing the Emotional Expressions of Non-Teammates ($N = 68$).

<table>
<thead>
<tr>
<th>Dummy-codes for team</th>
<th>Baseline Model</th>
<th>Individual Emotions</th>
<th>Type of Emotion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Amused  Happy  Afraid  Angry  Sad</td>
<td>Positive  Negative  Total</td>
<td></td>
</tr>
<tr>
<td>Control variables</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>−.01</td>
<td>.00</td>
<td>.01</td>
</tr>
<tr>
<td>White</td>
<td>.24*</td>
<td>.19</td>
<td>.25*</td>
</tr>
<tr>
<td>Age</td>
<td>.21*</td>
<td>.19</td>
<td>.20</td>
</tr>
<tr>
<td>General emotional accuracy</td>
<td>.19</td>
<td>.17</td>
<td>.25*</td>
</tr>
<tr>
<td>F-test of model</td>
<td>1.66*</td>
<td>1.70*</td>
<td>1.71*</td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>(16, 51)</td>
<td>(17, 50)</td>
<td>(17, 50)</td>
</tr>
</tbody>
</table>

Note: With the exception of model diagnostics, all coefficients listed are standardized $b$s.
- $p < .10$.
- $* p < .05$.
- $** p < .01$. 
Table 4. Correlations Between Team-Level Outcome Variables and Team Emotion Recognition Accuracy Corrected for Response Bias\(^a\) (\(N = 14\) teams).

<table>
<thead>
<tr>
<th>Outcome Composite</th>
<th>Positive Emotions</th>
<th>Negative Emotions</th>
<th>Total Accuracy</th>
<th>Regression Model with Positive and Negative Emotion Accuracy(^b)</th>
<th>Difference by Valence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Amused</td>
<td>Happy</td>
<td>Overall Positive</td>
<td>Afraid</td>
<td>Angry</td>
</tr>
<tr>
<td>Hackman #1:</td>
<td>.49*</td>
<td>.07</td>
<td>.33</td>
<td>-.27</td>
<td>-.25</td>
</tr>
<tr>
<td>Accomplished goal</td>
<td>.44</td>
<td>.38</td>
<td>.48*</td>
<td>-.48*</td>
<td>.07</td>
</tr>
<tr>
<td>Hackman #2:</td>
<td>.32</td>
<td>.05</td>
<td>.22</td>
<td>-.07</td>
<td>-.39</td>
</tr>
<tr>
<td>Work together again</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average liking among teammates</td>
<td>-.01</td>
<td>.19</td>
<td>.11</td>
<td>.14</td>
<td>-.66*</td>
</tr>
<tr>
<td>Member evaluation of team outcome</td>
<td>.43</td>
<td>.02</td>
<td>.26</td>
<td>-.14</td>
<td>-.26</td>
</tr>
<tr>
<td>Staff performance rating</td>
<td>.00</td>
<td>.17</td>
<td>.10</td>
<td>.44</td>
<td>-.53*</td>
</tr>
</tbody>
</table>

\(^a\)Response bias correction based on Wagner (1993).
\(^b\)Multiple regression model predicting the outcome variable from the total positive and negative emotion accuracy values. \(F\)-test degrees of freedom (2, 11).
\(^c\)Positive valence effect is the extent to which the team was more accurate in expressing positive rather than negative emotion. Values listed in this column are standardized linear regression coefficients from a model also controlling for the main effect of total team emotion recognition accuracy.
\(^d\)Values are \(Z\) coefficients testing the extent to which accuracy with positive emotion was a more positive predictor of outcomes than accuracy with negative emotion (Meng, Rosenthal, & Rubin, 1992). 
\(\sim p < .10\); two-tailed test.
\(* p < .05\); two-tailed test.
\(** p < .01\); two-tailed test.
aggregation of emotion recognition among the individuals who comprise the teams. The emergence of TERA, defined as the ability of individuals to understand their teammates’ emotions, was supported by data from a field study of intact teams. Taken together, positive and negative TERA accounted for 28.1% of the variance in team performance nearly a year later. By contrast, the GERA of these teams, akin to the group-level average of standardized emotion recognition skill examined in previous research, did not have the properties of a team-level construct. Thus, the study suggests that TERA is more than the sum of its parts. These findings provide support for an interpersonal perspective on the role of emotion communication in teams, beyond the exclusively individual-level focus driven largely by its early roots in the personality literature as well as the more recent interest in the umbrella concept of individual “emotional intelligence” (e.g., Goleman, 1995; Mayer et al., 2001; Rosenthal et al., 1979). The communication of emotion is an inherently social phenomenon, and the interactants are not simply interchangeable; we neither express emotion into a vacuum nor do we receive emotional messages from a vacuum. Thus, it is worthwhile to examine the communication of emotion beyond the simple aggregation of individual differences.

The current study also represents a methodological advance for research within organizational settings of communication via affective displays. In contrast with prior work on the recognition of context-free posed expressions, the present study used a well-validated protocol for emotion induction to elicit ecologically valid samples of emotional behaviors based on real events from past organizational settings. Constraints on the expression of emotion in organizational life make such communication subtle and complicated, guided by display rules (e.g., Ashforth & Humphrey, 1995; Fineman, 1993; Hochschild, 1983; Martin, Knopoff, & Beckman, 1998; Rafaeli & Sutton, 1989; Sutton, 1991; Van Maanen & Kunda, 1989), and the current study aimed to capture more of this subtlety and complexity. Further, such exercises elicited affective stimuli from expressors in participants’ actual workplace, rather than sampling from university students as in standardized tests.

**Antecedents of Team Emotional Accuracy**

After spending one week in team orientation, and not having begun in earnest their tasks together, randomly assigned work groups already varied significantly in the extent to which members could understand each other’s emotional expressions. What might have happened in team assignment or during that first week to explain these systematic differences?
Although the current study was designed to document these team differences rather than to posit and test a mechanism for them, prior research provides a foundation to speculate about their likely causes. First, some differences may have been present at the time of team formation, given that unacquainted pairs can vary idiosyncratically in the extent to which they can understand each other’s emotional expressions. Second, across the weeklong team orientation, workgroups may have differed in the extent to which team members grew acquainted and familiar with each other’s idiosyncratic styles. Group norms can emerge early in the life of teams (Flynn & Chatman, 2003), and teams may develop norms within even their first few minutes regarding the extent to which individuals share personal information (Polzer et al., 2002). Likewise, we argue that similar norms could quickly determine the extent to which members attend to, provide feedback regarding – and consequently learn about – each other’s emotional states. In our study, employees were significantly more accurate recognizing the emotional expressions of their own teammates than the expressions of members of other teams. Further, this accuracy at judging one’s own teammates varied meaningfully across the teams tested – whereas accuracy with non-teammates did not. These findings, taken together, appear to support a learning perspective on emotion recognition, in which individuals in some teams developed greater learning about each other’s emotional styles than those in other teams.

We speculate that teams high in team emotional accuracy may have shared more examples of their expressive style with each other, for example, due to greater time spent together, or due to deeper emotional exchanges during this time. Alternately, given that feedback about emotional expressions increases recognition accuracy (Elfenbein, 2006; Feldman, Philippot, & Custrini, 1992; Gillis, Bernieri, & Wooten, 1995), teams high in TEA may have been more explicit about their emotional states and thus provided more feedback that served to enhance learning. Further, members of some teams may have identified more closely with the team, and there is evidence that emotion recognition is more accurate when judging those with whom one identifies closely (Thibault, Bourgeois, & Hess, 2006).

One striking and unexpected finding in the study was that TERA for negative emotion predicted significantly lower task performance. The social functional perspective grounding this research emphasizes the adaptiveness
of understanding other’s internal states as a source of valuable information and feedback that can be used in order to adjust one’s behavior to be more appropriate for smooth interpersonal interactions. Why, then, would teams that achieve high accuracy in understanding each other’s negative emotions actually under-perform those teams that have less of this mutual understanding?

This result did not appear to stem from differences in the composition of the teams. Consistent with the teams’ random assignment, teams also did not vary systematically from each other in the extent of members’ levels of dispositional positive or negative affect, degree of self-monitoring, or their personality characteristics. Further, the first hypothesis of the study confirmed that the teams did not differ in the extent to which individual members were generally sensitive and accurate at understanding the emotions of others. Further, the results did not stem from response bias – whereby members of some teams would have simply endorsed positive choices more frequently, thus appearing to be more accurate – due to the signal detection methods used in the scoring, which control for the impact of such bias. It did not appear that good teams simply ignored the negative stimuli in their environment, given that Table 1 lists a positive correlation of $r = .30$ ($p < .02$) between accuracy at judging the positive and negative emotions of teammates. This indicates that teams more accurate with negative emotion were also more accurate with positive emotion, which is inconsistent with a response and attention bias explanation. Instead, given the positive correlation between TERA with positive and negative emotion, as well as the negative impact of negative TERA on performance and the trend of a positive impact of positive TERA on performance, it appears that this effect was driven by the non-overlapping variance. That is, those teams were poor performers who – controlling for their overall skill – were particularly diagnostic at understanding negative emotion rather than positive.

Why might this be the case? Through a process of emotional contagion (e.g., Barsade, 2002; Hatfield et al., 1994), understanding greater negativity from one’s colleagues may have invited reciprocity and thereby served to spread negative emotion throughout the team. The general benefits of positive affect in organizational settings, and deleterious consequences of negative affect, have been well established (e.g., Isen & Baron, 1991; Staw & Barsade, 1993; Staw et al., 1994). It may be that, in this particular team environment, negative emotion did not provide valuable information that others could use to adjust their workplace behaviors. According to the social functionalist perspective, what matters is not just the perception of an emotion, but most importantly what people do with the information once
they have perceived it. In this particular participant organization, jobs were highly structured, and teams operated within a larger structure characterized by command-and-control from senior staff members running a government program. The participants themselves were young adults generally in their first job. This suggests that the teams may not have had the skills, autonomy, and flexibility to change their behavior based on colleagues’ negative moods, and so such information may have served only as an aversive backdrop to the team’s work. Future research should test this moderator speculated to account for the negative finding – the degree of usability of the information that is gained by understanding the emotions of others.

**Study Applicability and Limitations**

Further work is necessary to develop a theory of emotional skills in teams, and additional empirical work is necessary to replicate and extend the current findings. In addition to exploring the antecedents and consequences outlined above, future work should seek to determine the boundary conditions of TERA. The current study took place within a single organization, which limits its generality. The participant organization was ideal in terms of random assignment of previously unacquainted individuals into teams, but represented a somewhat atypical workplace. Public service work is an environment where emotional skills may be particularly important and overburdened. However, the organization does represent a large and growing area within the United States, consisting of highly structured service-sector type employment undertaken by a relatively young and diverse workforce. The current study also does not provide a picture of how TERA might change over time. Given the inherently dynamic nature of communication and emotion, we can expect that TERA may evolve as team members become more closely acquainted with each other.

The present study has important limitations that future work would help to address. Most notably, in order to elicit naturalistic emotional expressions, the protocol made use of a well-validated procedure for the guided recall of emotion. However, there may have been individual differences in participant response to this procedure, and consequently some differences between participants’ actual workplace emotional expressions versus their enactment during the emotion interview. Even so, the lack of significant variation across teams along other individual difference variables suggests that such a factor would be unlikely to account for the current results. It would be valuable to
design future research to observe the accuracy of understanding emotions in
the workplace in real-time, using spontaneously occurring emotional expres-
sions. Further, future research may separate out the full-channel displays
that simultaneously include facial expressions, paralinguistic cues, body
movement, and implicit verbal content in order to help determine which cues
were the most influential for participant judgments.

CONCLUSION AND IMPLICATIONS

Emotional sensitivity has long been viewed in the research literature as an
individual intelligence or stable personality trait, which is consistent across
interaction partners. This study supported, instead, a perspective on emo-
tional skill that may be fruitfully conceptualized beyond the individual level.
Past research demonstrated that work groups could meaningfully share
emotional states. The current results show that teams can meaningfully
share the ability to perceive each other’s emotional states – and this ability
appears to be more than a mere aggregation of individual differences in
general skill. Future research has much to gain by seeing the effectiveness of
emotional skill as a property of teams. As current trends continue towards
increasing teamwork and interdependence, the need for teams to improve
the quality of their coordination is likely to become more critical. TERA
may be one of the skills necessary for such coordination. Managers may
wish to consider interpersonal dynamics early in the life of teams, and to
include exercises and activities designed to enhance the mutual understand-
ing of emotional states.

ACKNOWLEDGMENTS

We thank Teresa Amabile, Robin Ely, and Charles O’Reilly for their de-
tailed comments and support. We are indebted to David Kenny for his
statistical advice. Additionally, we are grateful for advice and assistance
from Amy Edmondson, Dan Elfenbein, Eena Khalil, Sei Jin Ko, Eric Lauer,
Debi La Plante, Ying Liu, Steve Nowicki, and Justin Saif. Preparation of
this article was supported by a National Science Foundation Graduate
Student Fellowship and a Harvard Business School dissertation research grant.
REFERENCES


