

Identifiability, Performance Feedback and the Köhler Effect

Norbert L. Kerr, Lawrence A. Messé, Ernest S. Park, Eric J. Sambolec

► **To cite this version:**

Norbert L. Kerr, Lawrence A. Messé, Ernest S. Park, Eric J. Sambolec. Identifiability, Performance Feedback and the Köhler Effect. Group Processes

Intergroup Relations, 2005, 8 (4), pp.375-390. <10.1177/1368430205056466>. <hal-00571613>

HAL Id: hal-00571613

<https://hal.archives-ouvertes.fr/hal-00571613>

Submitted on 1 Mar 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Identifiability, Performance Feedback and the Köhler Effect

Norbert L. Kerr and Lawrence A. Messé

Michigan State University

Ernest S. Park

North Dakota State University

Eric J. Sambolec

Michigan State University

Research, starting with Köhler (1926), has demonstrated a type of group motivation gain, wherein the less capable member of a dyad working conjunctively at a persistence task works harder than comparable individuals. To explore possible boundary conditions of this effect, the current experiment systematically varied the amount and timing of performance feedback group members received. Results showed: (a) continuous feedback of both members' performance was not necessary for producing the effect; (b) the effect was attenuated, but not eliminated by delaying and restricting feedback, such that group members only learned which of them was the less capable worker (but not how long s/he persisted) sometime after the task trial was completed; and (c) the motivation gain was eliminated in the absence of any performance feedback (i.e. when neither worker could tell who quit first nor how long either had persisted). Some implications of these results for currently viable explanations of the Köhler effect are discussed. It is concluded that the effect is likely to result from several distinct processes.

KEYWORDS feedback, groups, motivation gain, performance

CONSIDERABLE research has documented that members of task performing groups often engage in *social loafing*—that is, they exert less effort than comparable individual performers (see Karau & Williams, 1993, and Shepperd, 1993, for reviews). Demonstrations of the opposite result, where one or more workers try harder when performing in a group or team than as individuals have been less frequent. Yet there is some evidence of such group motivation gains in the literature, particularly in the last decade or so (e.g. Hertel, Kerr, & Messé,

2000; Stroebe, Diehl, & Abakoumkin, 1996; Williams & Karau, 1991). The present study builds on this research by exploring boundary conditions of one such group motivation-gain phenomenon—the tendency of less able workers to perform better when members of a

Author's note

Address correspondence to Norbert L. Kerr, 430 Baker Hall, Michigan State University, East Lansing, MI 48824-1117, USA [email: kerr@msu.edu]

team working under conjunctive task demands (where the performance of the least capable member defines the group's level of performance; Kelley et al., 2003, Entry #9; Steiner, 1972) than when working individually. This phenomenon has been named after Otto Köhler, who first reported evidence for it nearly 80 years ago (Köhler, 1926, 1927). While more recent investigations (e.g. Hertel, Deter, & Konradt, 2003; Hertel, Kerr, & Messé, 2000; Messé, Hertel, Kerr, Lount, & Park, 2002; Stroebe et al., 1996) have replicated the Köhler motivation gain effect, none have identified its boundary conditions.

Identifying and explaining such motivation gain phenomena is valuable for several reasons. As strictly an applied matter, much important work is done by groups, increasingly so in organizational teams over the last few decades (Guzzo & Dickson, 1996; Hackman, 2002). Identifying conditions that will maximize group member task motivation holds promise for enhancing the productivity of such groups. Theoretically, Allport (1962) has suggested that clarifying the relationship between individual and group behavior is the master problem of social psychology. The voluminous literature on social loafing has led to the misconception that group work conditions are inherently or generally demotivating. Research on the Köhler effect and other motivation gains, such as the *social compensation* effect (Williams & Karau, 1991), has belied that misconception. More importantly, learning more about the motivating properties of groups could contribute to developing a more powerful, unified theory that can account for both group motivation losses and gains. For example, Karau & Williams (1993) have suggested that instrumentality \times value theory (e.g. Vroom, 1964) may well serve such a function.

Some research suggests that the Köhler effect may result from the weaker member at a conjunctive group task feeling particularly indispensable to the group's success. For example, in Hertel et al. (2000, Experiment 2) participants worked in two-person teams on a physical persistence task, but they did so under different task demands—either conjunctive or additive, where the group score

was simply the sum of individual scores. These conditions differed markedly with regard to indispensability, given that the weaker worker's performance was substantially more important to the team's outcome under conjunctive task demands. Consistent with the indispensability explanation, results revealed a significantly greater motivation gain under conjunctive task demands than under additive demands. Other research suggests that the effect could result at least in part from the weaker member's social comparison with the stronger member and, as a consequence, upwardly revising his/her performance goal and/or feeling heightened implicit competitiveness (Kerr, Messé, Seok, Sambolec, & Park, 2005; Lount, Messé, & Kerr, 2000; Stroebe et al., 1996). For example, when a weaker coactor can observe the performance of a more capable coactor, reliable motivation gains have been observed, although they tend to be weaker than those observed under conjunctive group demands (Kerr et al., 2005; cf. Hertel et al., 2000, Experiment 2). Finally, still other research demonstrates that the identity of one's partner may moderate the effect, suggesting that impression management concerns may contribute to it. For example, Lount et al. (2000) found that males exhibited a larger Köhler effect when working with a more capable female than male partner.

One approach to resolving such theoretical puzzles is to conduct competitive experimental tests of the alternative theories (e.g. Hertel et al., 2000, Experiment 2; Kerr et al., 2005). Yet another approach is to empirically identify factors or conditions that moderate an effect, and then evaluate and/or revise contending explanations in the light of the discovery of those moderators. The research demonstrating the moderating role of the discrepancy of member abilities (Messé et al., 2002) is a good illustration of the latter approach. We adopted this 'identify moderators' approach in the current study by examining the impact of various amounts and timing of performance feedback on the magnitude of the Köhler effect. We chose to focus on aspects of workers' performance feedback, in part, because

previous research demonstrated that one form of performance feedback—knowledge of relative ability in the group—plays an important role in determining how much harder the weaker member tries under conjunctive conditions (Köhler, 1926; Messé et al., 2002). In addition, such a manipulation also permits tests of some of the current explanations for the effect.

Performance feedback and identifiability

All research investigating the Köhler effect has employed continuous feedback of how well all group members are doing, and publicly identified the member who quits first. For example, with the arm-lifting persistence task paradigm that we have used in most of our prior work, dyad members performed in sight of one another. They could readily monitor each other's performance and notice who was becoming more fatigued. Moreover, when one member's arm dropped sufficiently to touch the horizontal trip wire, end the trial, and thereby determine the group's score, it was evident to both who this weaker member was and how long s/he had persisted. The present study asks whether such continuous performance feedback and identifiability are necessary for or facilitative of the Köhler effect. Answering these questions is not only of obvious interest for establishing the effect's generality, but could also shed light on its likely cause(s). Within the three general causal explanations that have been advanced—indispensability, social comparison, impression management—there are a number of possible subcases or versions which make different assumptions about the underlying processes and can be competitively tested via systematic variation in the nature of performance feedback and member identifiability. For example,

- As noted above, there is evidence that the effect may be driven by the *indispensability* of effort by the group's weakest member at a conjunctive task (Hertel et al., 2000; cf. Kerr,

1990; Kerr & Bruun, 1983). There is, though, more than one way in which this indispensability mechanism might work. It could be that the simple knowledge that the task is conjunctive and hence, that the group's score will be determined by the poorest single member performance is sufficient to motivate supranormal effort by all group members. We call this version of the indispensability explanation—which stresses the mere possibility of being indispensable to the group—the *possible indispensability* version.

- Alternatively, it could be that before a worker increases her/his effort, s/he also needs some information indicating not just that it is possible, but that it is probable that s/he is indeed that least capable member (a *probable indispensability* version of the general indispensability explanation).
- The *goal setting* version of the social comparison explanation, advanced by Stroebe et al. (1996), suggests that the less capable member must already possess the knowledge that his/her partner is more capable (and, perhaps, of how much better his/her partner is) before any motivation gain would occur.
- However, an *implicit competition* version of the social comparison explanation does not require this prior knowledge. That is, if it is the concern that one might be outperformed by a fellow group member that underlies the effect, all that may be required to boost one's motivation is the belief that one could be revealed as inferior to one's partner.
- Finally, any *impression management* explanation—based on striving to create a favorable impression—would predict little or no motivation gain under conditions where one's performance could not be publicly identified.

The present design permitted an empirical check on these and other interesting possibilities by systematically varying the nature of the performance feedback and identifiability of the poorer performer. There were four experimental conditions. (1) Our Full-feedback condition paralleled the prior paradigms—group members had continuous feedback of one

another's performance and were identifiable to one another. All of the alternative explanations presented above predict a robust Köhler effect in this condition (see Table 1). The three remaining experimental conditions increasingly restricted performance feedback information. (2) In an End-of-trial feedback condition there was no feedback during the trial of how well each group member was doing. Only at the end of the trial did the feedback indicate who had quit first. Although each group member had some information about partner ability as a trial commenced (viz. he knew that his partner was still persisting, and hence, had not yet been revealed to be the weaker member), he had no explicit information to indicate who ultimately would prove to be the weaker member. This condition was included primarily to determine if joint, continuous feedback of group members' performance is necessary for the Köhler effect; none of the theoretical explanations presume that it is, and so again, all predict that the effect will be replicated in the End-of-trial condition (see Table 1). (3) More theoretically interesting is the Delayed-feedback condition. Here information about who quit first (but not when) was provided, but only some time after the trial was over. This condition created an interesting type of conjunctive task in which a worker has no way of knowing while engaged in the task whether or not his partner has already quit and thereby ended the trial. Thus, one cannot tell if one's own continuing efforts on the group's behalf may be wasted (i.e. if one can make any unique contribution to the group; Harkins &

Petty, 1982; Weldon & Gargano, 1988). The latter possibility seems likely to be demotivating and could well attenuate the Köhler effect. Furthermore, in this condition, two of the theoretical explanations predict no Köhler effect (see Table 1). The *probable indispensability* interpretation holds that one needs information during the trial suggesting that one's efforts are indispensable to the group, but such information is unavailable in this condition until after the trial is over. And the *goal setting* explanation requires *prior* knowledge of one's partner's better performance before one sets and strives toward a higher goal. Again, no such information is available until after the trial is completed in this condition. Two other explanations predict a motivation gain, but suggest that it could be somewhat attenuated (see Table 1). The Delayed-feedback condition offers no chance for intragroup comparison *during* the performance trial, but does *after* some delay; to the degree that one strives to avoid an invidious future comparison as well as an immediate one, the *implicit competition* mechanism would predict some motivation gain. Likewise, the delayed feedback of this condition only reveals who quit first in the group, but nothing about how long he persisted. So even if one does poorly, less negative information will be revealed to others. Thus, the *impression management* explanation also predicts an attenuated effect. (4) Finally, in the No-feedback condition, no one could tell either who quit first or how long either had persisted. Only one explanation still predicts a significant motivation gain in this condition; the conjunctive task

Table 1. Predictions of alternative theoretical explanations

Theoretical explanations	Experimental condition			
	Full-feedback	End-of-trial feedback	Delayed-feedback	No-feedback
Possible indispensability	++	++	++	++
Probable indispensability	++	++	0	0
Goal setting	++	++	0	0
Implicit competition	++	++	+	0
Impression management	++	++	+	0

Note: ++ signifies a robust Köhler effect; + signifies an attenuated effect; 0 signifies no effect.

demands are known to all, so if only the possibility that one's efforts might be indispensable to the group is sufficient to produce the effect, the *possible indispensability* explanation predicts that the effect should be replicated here (see Table 1). However, the impoverished feedback provided in this condition eliminates information held to be necessary for the effect by all the remaining explanations (viz. information that one's performance is more likely than one's partner's to be particularly indispensable to the group; information that one's partner is setting a higher performance standard; information indicating that one is or may be outperformed by one's partner; or performance information that could contribute to a negative impression by one's partner). Although this design does not provide direct competitive tests between all possible theoretical models, it can narrow the range of plausible explanations (Platt, 1964), as well as establishing possible boundary conditions for the effect.

Method

Participants

A total of 217 male undergraduates at Michigan State University participated in this research. All participants received credit in their introductory psychology course. In addition, all were made aware that they could earn a monetary reward for performing well.

Experimental design

There were four experimental dyad conditions and an individual control condition. In the latter, subjects performed all trials in isolation and individually. This control condition was included to estimate and then take into account the effects of fatigue and boredom on effort. The four experimental conditions were, as described above: Full-feedback, End-of-trial feedback, Delayed-feedback, and No-feedback.

Procedure

The experimenter greeted each participant at a designated location away from the research laboratory. The participant was told that he would be working alongside another person,

who had arrived a few minutes earlier and was already set up in a lab booth. He was also informed that the experiment was automated and the experimenter would not be in the lab room during the session. The participant was then asked to read and sign a consent form and follow the instructions at the bottom of the page. At this point the experimenter left, ostensibly to go to his/her office. In fact, the experimenter went to the lab and entered a curtained booth (located next to the participant's booth) which contained a VCR with instruction tapes and the controls for all experimental equipment.

The instructions at the bottom of the consent form directed the participant to go to the lab room, seat himself in the available booth, and don a set of headphones. (The other, already-closed booth, ostensibly was occupied by the second participant, but in reality the experimenter sat there.) Prerecorded verbal instructions were delivered over the headphones, which also served to mask sounds from the neighboring booth.

The instructions directed participants to open up the booth curtains so that they could see a television monitor on which further directions and information would be presented. These instructions stated that the purpose of the study was to examine the persistence of people working at a physical task—viz. attaching a weighted bracelet to one's arm and holding it extended horizontally at shoulder height for as long as possible. The video showed how to secure the weighted (1.8 kg) bracelet around one's wrist with Velcro straps.¹

Participants were told that they would be performing several trials during their session, but not the exact number (which was four). At the beginning of each trial, the videotape would announce, 'Get Ready'. At this command, participants were to place their arm (with the weighted bracelet attached) straight out, with their elbows locked, aligning their wrist to a mark on the wall of the booth. This mark was approximately 12 inches above an infrared trip beam. A moment later, the 'Go' signal would be given, marking the start of the trial. Participants were to hold their arm above the trip beam for as long as they could without risking undue

fatigue or injury. When a participant's arm dropped to the horizontal, the beam would be broken and the trial would be over. The trial score, automatically recorded by computer, was simply the number of seconds that the participant persisted at the task. Inside each participant's booth was a signal light that illuminated when the beam was broken. At this point, the participant was to rest his arm and wait for the next trial to begin.

To further motivate the participants, all participants were told before they began the first trial that they had an opportunity to earn money through their performance at the persistence task. The experimenters would sum up all of the scores on all of the trials that he and the other participant completed. They were then told that when data collection for the study was finished, the experimenters would randomly pick one of the pairs taking part in the study. The chosen pair would receive five cents for every second that the pair of participants lasted at the task, with the total sum being split evenly between the two participants. Hence, there was an extrinsic incentive that was the same for all conditions, including the individual (coacting) controls.

Every participant first completed one trial with his dominant arm, and then a second trial with his non-dominant arm. For these first two trials the participant could not see how the person in the other booth was performing. Moreover background noise that the participant heard through his headphones masked any possible auditory clues to the other's activities. The videotape was recorded so that each trial could last for a maximum of 4.5 minutes. If a participant persisted this long (a very rare event), he was instructed to lower his arm and prepare for the next trial.

The experimental manipulations were imposed between the second and third trials, as follows:

Individual control condition. After a wait sufficient to give as much rest time to these participants as to those in the other, experimental conditions (who received additional instructions), participants in this control condition followed the same procedure as before. They

completed two more trials, the third with their dominant arm and the fourth with their non-dominant arm. Note that—except for the differences between individual and group performance of direct interest here (viz. in collaborative work and its attendant differences in indispensability of effort, mutual performance feedback and social comparison opportunities)—the work conditions for these individual controls were the same as for the experimental dyad conditions. In particular, there were no experimental-control differences in 'mere presence' of others (Zajonc, 1965) or in possible distracters.

Experimental dyad conditions. After Trial 2, participants in the experimental conditions were told that they would henceforth work at the persistence task with the other person in a two-person team. On these dyad trials, the team score would be determined by the first team member to hit the beam; thus, the task demands were conjunctive. The two persons to be randomly selected to receive a cash reward for their performance would be on the same team and the reward for dyad trials would be based on team performance.² Like their individual control counterparts, the participants assigned to work as dyad members performed the third trial with their dominant arm and the fourth with their non-dominant arm. There were four different dyad conditions.

(1) Full-feedback Condition: In this condition, participants were able to continuously monitor one another's performance during each dyadic trial. Prior to each trial, each extended a collapsible radio antenna that was attached to the weight bracelet. The extended antennas could be seen from the neighboring booths, so that each could always see the position of coworker's arm by the height and angle of that person's antenna. In addition, there was a heretofore-dormant stand set up in front of the television monitor. There were two lights on the stand, one labeled 'A' (the ostensible coworker's booth) and the other 'B' (the actual participant's booth). When either member's arm broke his infrared beam, his respective light on the stand would illuminate, thus signaling the end of the trial to both.

Therefore, participants would be immediately aware of which member quit first for each dyad trial, and when. However, the computer would only record the team score and the experimenter, who was ostensibly not in the room, could not determine who quit first.

During the third and fourth (dyad) trials, the experimenter simulated the performance of a partner, who always persisted longer than the actual participant (that is, the experimenter's arm stayed well above the beam until the participant's arm broke it; this was not difficult because the experimenter did not wear a weighted bracelet). After the third trial, the light on the stand, which showed that the actual participant had quit first, remained illuminated until just before Trial 4 began. And as with the earlier trials, between trials the participant rested for a few minutes, moved the bracelet to the other arm, and waited for the next trial to begin. After Trial 4, the signal light, once again indicating that the actual participant had quit first, remained on until the end of the session.

(2) End-of-trial feedback Condition: The instructions and procedure for this condition were identical to those in the preceding, Full-feedback condition with one exception: There was no mention of the antennas, which remained collapsed throughout the entire session. Thus, the participant could not monitor how well his partner was doing during either dyad trial, and only discovered at the end of the trial (when his arm hit the beam and both his in-booth and the signal-stand lights lit up) that he, himself, had quit first, and hence, ended the trial.

(3) Delayed-feedback Condition: The instructions and procedure for this condition were identical to those in the preceding, End-of-trial condition, with one exception: When the participant's arm hit the beam, his in-booth light illuminated immediately, but the light on the external stand (which was visible to both dyad members) did not become lit until a few seconds before the start of the next trial. Hence, the feedback about which dyad member quit first was delayed for a few minutes after the end of the trial. Moreover, this feedback did not reveal how long each member

had worked. While he was working, the participant could not tell if he or his partner had quit first, but he was aware that this information would be revealed to both of them just before the start of the next trial.

(4) No-feedback Condition: The instructions and procedure for this condition were identical to those in the preceding, Delayed-feedback condition, with one exception: there was no mention or use of either the antenna or of the external light stand. Thus, participants were aware that neither dyad member (nor the experimenter) would know which member quit first on any trial. Participants in this condition were informed that because of the way in which the equipment was wired, the computer would record and save only the length of the trial, and not which person ended the trial. They were also instructed that the light in their booth would come on as soon as they had broken the beam. However, they would never know which member of the team had quit first.

Results

Analysis of control individuals' performance data

A preliminary analysis of the performance of the 38 participants in the Individual Control condition was conducted via a 2 (Trial block) \times 2 (Trial within block) repeated-measures analysis of variance (ANOVA) on the amount of time per trial that they kept their arm above the beam. The latter factor is functionally an Arm (dominant vs. non-dominant) variable, since participants performed Trials 1 and 3 with their dominant arms and Trials 2 and 4 with their non-dominant arms. There was a significant Trial (Arm) main effect ($F(1,37) = 4.39, p < .05, \eta^2 = .11$) indicating that control participants persisted longer with their stronger, dominant arms (174.8s, $s = 37.40$) than their non-dominant arms (163.8s, $s = 48.79$). More importantly, there was a significant Block main effect ($F(1,37) = 43.83, p < .001, \eta^2 = .54$), which confirmed a strong fatigue/boredom effect. Specifically, participants persisted longer on their first trials with each arm (Block 1 mean = 189.0s, $s = 44.50$) than on their second attempts

(Block 2 mean = 149.6, $s = 44.05$). The Block \times Trial (Arm) interaction was not significant ($F(1,37) = .88$). Hence, the fatigue effect was about the same for the dominant and non-dominant arms; the average control subject persisted 39.4s less on the second testing of an arm. This value established a potential productivity baseline (Steiner, 1972) for our dyads. Thus, any dyad participants who showed less than a 39.4s decline in their persistence from Block 1 (when they were working individually) to Block 2 (when they were working in the dyad) manifested a motivation gain.

Analysis of dyad members' performance data

The primary data were the participants' change scores from Block 1 to Block 2. There were two such scores, one for the dominant arm (Trial 3 – Trial 1) and the other for the non-dominant arm (Trial 4 – Trial 2). In order to adjust for the constant fatigue factor estimated from the Individual Control participants' performance, the fatigue correction (39.4s) was added to each of these change scores. This provided a personally normalized estimate of the magnitude of motivation gain for the dominant arm ($MG_{\text{dom}} = \text{Trial 3} - \text{Trial 1} + 39.4\text{s}$) and for the non-dominant arm ($MG_{\text{nondom}} = \text{Trial 4} - \text{Trial 2} + 39.4\text{s}$); positive values indicated motivation gains, while negative values indicated motivation losses.

A 4 (Feedback conditions) \times 2 (Arm: dominant vs. non-dominant) ANOVA with repeated measures on the latter factor was performed on these motivation gain scores. Only the Feedback main effect was significant ($F(3,162) = 12.45$, $p < .001$, $\eta^2 = .19$). There was no indication that participants had greater motivation gains with one arm than the other (Arm main effect $F(1,162) = .09$). Further, the effects of feedback condition on motivation gains were not moderated by Arm (Feedback \times Arm interaction $F(3,162) = .44$). Hence, below we report results for the average of the motivation gain scores across the two arms.

The mean motivation gain scores for the four experimental conditions (with 95% confidence intervals) are plotted in Figure 1. In the following presentation, we consider the conditions

starting with those providing the most information to group members and moving to those providing less information. First, consider the Full-feedback condition which mirrors the conditions used in prior studies of the Köhler effect. Dyad members have continuous feedback of one another's performance and the identity of the first person to quit is known to both. This condition replicated the robust Köhler effect observed in the prior studies. The participants worked 41.7s ($s = 36.36$, $n = 41$) longer than the individual controls; this value is significantly greater than zero. The same conclusion results from including the individual condition as a control and doing Dunnett tests comparing experimental conditions with it. The Full-feedback condition produces a highly significant motivation gain ($p < .001$).

In the End-of-trial feedback condition dyad members had no continuous feedback of each other's performance. The participant in this condition—like the participant in the Full-feedback condition—is functionally (by experimental contrivance) the less capable member and, until the moment he himself quits, is continually aware that his partner has not yet quit. However, unlike in the Full-feedback condition, he has no other information indicating that he is, indeed, the less capable and more indispensable member of the pair (e.g. seeing that his own arm is starting to drop with fatigue whereas his partner's is not). In this condition, we found that the motivation gain was only slightly lower (mean = 35.01s, $s = 37.25$, $n = 39$) than in the Full-feedback condition. Both the 95% confidence intervals plotted in Figure 1 and a Dunnett test indicated that the motivation gain effect in this condition is highly reliable ($p < .001$). More importantly, it did not differ significantly from the gain observed in the Full-feedback condition (based on a post hoc, Newman-Keuls comparison). Thus, within the limits of statistical inference, these data suggest that continuous performance feedback is neither necessary nor facilitative for the Köhler effect.

Next, we considered the Delayed-feedback condition. Here, the performance feedback was delayed such that coworkers only learn some

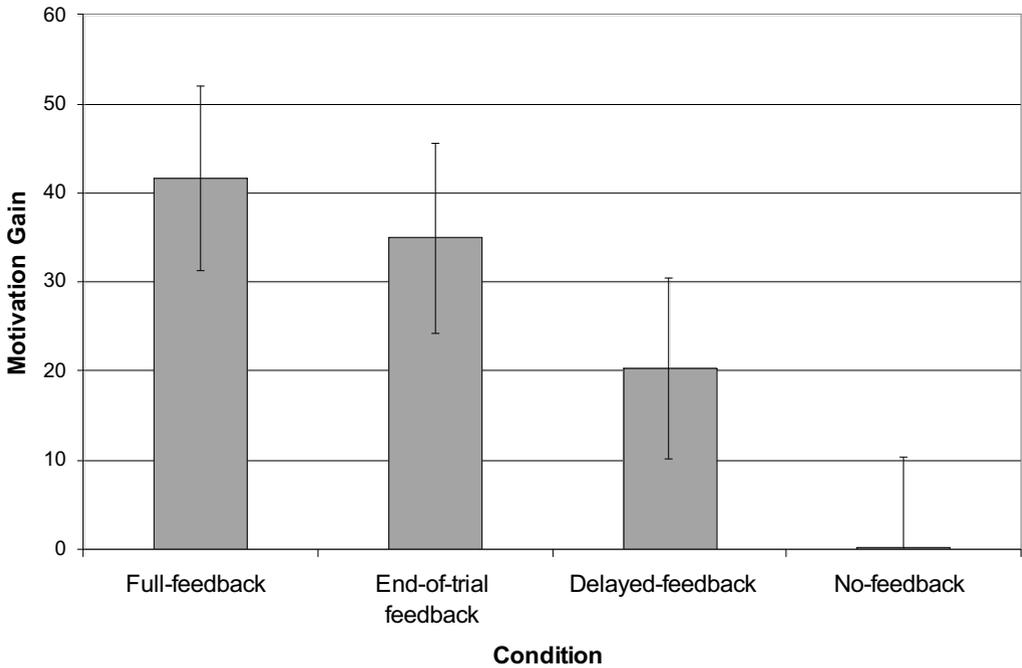


Figure 1. Motivation gains in performance feedback conditions.

time after the trial was completed which member quit first, but not when. In this condition, there was a significant ($p < .03$, Dunnett test) motivation gain (mean = 20.30s, $s = 30.21$, $n = 43$). However, post hoc Newman-Keuls comparisons revealed that its magnitude was significantly less than that in either the Full-feedback or the End-of-trial feedback conditions. The reduction in feedback for this condition clearly attenuated the Köhler effect, but did not eliminate it.

Finally, consider the No-feedback condition. Here, there was no information available to indicate which dyad member quit first. Figure 1 reveals that, in fact, the mean performance in the No-feedback condition (mean = .21s, $s = 30.70$, $n = 43$) was considerably (and significantly, according to Newman-Keuls post hoc tests) smaller than the already attenuated gain observed in the Delayed-feedback condition. Moreover, in the present No-feedback condition, there was no significant motivation gain at all ($p > .99$, Dunnett test).

Discussion

In this study we explored the effects of varying the amount and timing of information group members receive about one another's performance while working at a conjunctive persistence task. Our specific goals were to identify possible feedback-based boundary and facilitative conditions for the Köhler motivation gain effect, and thereby, to inform the search for the underlying cause of this motivation gain phenomenon (and, perhaps, other such phenomena yet to be discovered). Our findings, summarized in Figure 1, justify a number of conclusions.

Performance feedback and the generality of the Köhler motivation gain effect

Our results suggest that continuous feedback of the absolute (and hence, relative) levels of performance of all group members is not necessary for the Köhler effect—the magnitude of the effect when group members had this information (Full-feedback condition) was not

significantly greater than when group members did not have it (End-of-trial feedback). This lack of difference has a number of interesting implications: (a) the external validity of the Köhler effect is extended to tasks without such continuous performance feedback, and, in addition, (b) the phenomenon can be studied using simpler experimental paradigms that dispense with such feedback (cf. Hertel et al., 2003).

At the other extreme—of no performance feedback whatsoever and complete anonymity of group member contributions—we found no trace of the Köhler motivation gain effect. Thus, to produce this phenomenon, it is not sufficient that the group's tasks simply have conjunctive task demands; in addition, it appears that there must be some possibility of group members (or perhaps, others; Harkins, 2001) knowing or being able to identify the relative abilities of group members. Besides eliminating a number of group performance contexts as likely settings for producing this motivation gain, any feasible explanation for the effect must incorporate this boundary condition.

Finally, we identified some conditions that attenuated but did not eliminate the effect. Those conditions were (a) providing no information to group members about how well they each were performing as they worked, but (b) providing information about which member performed less well sometime after the task was completed. Clearly, for those seeking to maximize member motivation and group performance at conjunctive tasks, it would be useful not to delay relative performance feedback.

It is important to note that, even with the limiting and boundary conditions discovered here, the set of group tasks for which we might reasonably expect to obtain the Köhler motivation gain effect is potentially still quite large (Kelley et al., 2003). The full set of group tasks with conjunctive features includes:

- Unitary persistence tasks (such as the group task used in this experiment) or other performance tasks in which group members simultaneously engage in the identical task

and task demands limit group performance to the poorest performance in the group. The conjunctive nature of the task can stem from the rules of the task and the definition of collective performance (as in our experimental task). But they can also stem from group members being yoked or tethered, so that a slow or poorly performing member holds back the rest of the group (e.g. tethered mountain-climbing teams or a group of commuters trying to get home along a single-lane highway).

- Divisible tasks where the complete task is not finished or satisfactory unless each group member has finished his/her portion satisfactorily. An example would be a group of four coauthors who divided the task of writing a research article, with one doing the Introduction, another the Method, etc. An inadequate job by any one of the coauthors would result in an inadequate combined draft. Hertel et al. (2003) have reported a robust Köhler effect for just such a divisible task. Dyads worked at a business simulation in which they matched customers' computer and computer-furniture needs with a list of product options, trying to find the package that provided the highest profit for the firm. The order was not complete until both parts were complete (i.e. filling the order for computer components and filling the order for computer furniture). Compared to control individuals responsible only for completing their own orders, dyad members responsible for the computer orders worked faster and more accurately with a partner who worked faster than they did. It is also worth noting that this conceptual replication of the Köhler effect was obtained for a fairly complex, cognitive performance task (rather than the fairly simple, motor persistence task we used in the present and several other experiments).
- Sequential group tasks, for which one group member cannot begin his/her job until preceding members have completed their jobs. An illustration might be a swimming relay team. Each member must wait for all the preceding members to complete their

laps before s/he can enter the pool. A sub-standard performance by any member ensures a poor team performance.

To be sure, the present study suggests that there does need to be some level of publicness and timely feedback about who can do what before we would expect to see a robust Köhler effect in such tasks. However, these limiting conditions still leave a very large and interesting set of collective tasks where we might expect reliable Köhler motivation gains. Of course, the external validity of the effect is largely an open question and needs to be explored empirically in future research.

Narrowing the range of plausible theoretical explanations

There are currently a number of plausible explanations for the Köhler motivation gain effect. Although we have not completed the search for the 'correct' explanation(s) in this study, we have made some progress toward that goal. In particular, the present data refute at least three variants of older, general explanations and suggest a novel variant.

'Indispensability of effort' as an explanation

One explanation for the Köhler motivation gain (Kerr, 1990; Kerr & Bruun, 1983) attributes it to the fact that when a group faces a conjunctive task, the performance of the group depends vitally on the performance of one member—the least capable member—and it is the indispensability of that member's efforts that prompts him/her to work especially hard. The absence of any motivation gain in our No-feedback condition contradicts one version of this indispensability explanation—one that assumes that mere knowledge that the task is conjunctive is sufficient to produce a motivation gain (the *possible indispensability* version). More than this knowledge is clearly required.

There are other versions of the indispensability explanation, however. It could be that before a worker increases her/his effort, s/he also needs some information indicating that it is probable that s/he is indeed that least capable member (the *probable indispensability*

version). Just such information was available in both our Full- and End-of-trial feedback conditions, which produced the strongest motivation gains. In both of these conditions, there was a type of continuous feedback—information confirming that one's partner was still persisting. In the Full-feedback condition, there was clear visual evidence that the partner never seemed to tire. In the End-of-trial condition, the information was less explicit and detailed, but informative nonetheless. As the subject began to tire and to approach his own limit of endurance, the fact that his partner had not yet reached his/her limit made it increasingly likely that the subject was the weaker of the pair. Thus, both conditions provided information to the subject during the trial and prior to his own decision to quit that suggested that he was very probably the weaker member of the dyad.

This explanation would suggest that when such information was missing, one would observe no motivation gain. Hence, it is consistent with the null effect in the No-feedback condition, where one had no means of ascertaining that one was the less capable member. However, it would also predict a null effect in the Delayed-feedback condition. Here, only some time after the trial was completed could one learn that one was the less capable member of the dyad. Nevertheless, there was a significant (although attenuated) motivation gain found in this condition. Therefore, the *probable dispensability* explanation, by itself, cannot fully explain the overall pattern of results.

Participants in the Delayed-feedback condition should have had no information to suggest to them or to their partners that their efforts were indispensable to the group's success. However, group success/failure might not be the only outcome of concern. Being publicly revealed to be the 'weak link' in the group (and, hence, responsible for a poor group performance, having let one's partner down, or having limited the effectiveness of one's partner's performance) may also be viewed as an aversive outcome. Increased effort would be instrumental in avoiding such an outcome in all but the No-feedback condition.

These two concerns—with the indispensability of one's efforts for the group's success and with being revealed as the team's 'weak link'—together could account for the pattern of motivation gain differences found across all four dyad conditions.

Social comparison explanations As noted earlier, Stroebe et al. (1996) have argued that the Köhler motivation gain may result from a process of social comparison in which learning that another is more capable leads the less capable group member to raise his/her performance goals. This goal setting explanation cannot account for the significant (attenuated) motivation gain observed in the Delayed-feedback condition, where the participant could, at best, only compare with his partner sometime *after* the task was completed. An advocate of this theory might argue that a participant could compare himself with his partner when he received the delayed Trial 3 feedback and that the attenuated motivation gain observed in this condition was the result of a strong gain on the subsequent Trial 4, averaged with no gain on Trial 3. The nonsignificance of the Feedback condition \times Trial interaction contradicts this argument, but to give it its strongest possible test, we specifically examined the effect of Trial (i.e. Trial 3 vs. Trial 4) within the Delayed-feedback condition. The means were in the direction suggested by this argument (viz. Trial 3 mean = 16.7s motivation gain, Trial 4 mean = 24.04s motivation gain), but the difference was clearly nonsignificant ($t(42) = .91, p = .37$). Hence, the goal setting explanation also cannot account for the full pattern of results from this study.

But perhaps it is only the 'goal setting' notion that is problematic. Maybe group members do engage in social comparison, but the result of that process is (or can be) something other than setting a personal performance goal. One possibility, hinted at by Stroebe et al. (1996), is that the 'something' might be an attempt to successfully compete with the other—to outperform or, perhaps, to avoid being outperformed. Such an *implicit competition* explanation is fairly consistent with the full pattern of the

present results. When group members believed that their partner was more capable (as in the Full-feedback or End-of-trial conditions), they increased their effort to successfully compete with him. When there was no possibility of either group member successfully outperforming the other (as in the No-feedback condition, where no one could tell who performed better), there was no motivation gain. The Delayed-feedback condition represents an interesting intermediate case in which one had little idea of which was the stronger group member as one performed, but one might later be shown to have been outperformed by one's partner. If the likelihood of working harder declines with the certainty that one is about to be outperformed, then we would expect conditions that reduced this certainty (like the Delayed-feedback conditions) to attenuate the motivation gain effect. This is exactly what happened.

Another possibility is that the 'something' underlying a social comparison process is an attempt to evaluate and validate one's own task ability, basic premises of the original social comparison theory (Festinger, 1954).³ That is, one may want both to see just what one is capable of (to evaluate one's ability) and to meet (or surpass) available social comparison targets (if possible, to show that one has relatively high ability). Such a *self-evaluation* version of the more general social comparison explanation predicts high striving when the 'other' is likely to be more capable (Full-feedback and End-of-trial conditions), and no increase in striving when there is no way of telling just how the group members compare (No-feedback condition). Again, the Delayed-feedback condition presents a less immediate or certain opportunity for self evaluation. In this condition, one's performance could eventually be compared with another person's and could be shown to be inferior. However, the risk of these outcomes is only a possibility in this condition, rather than a high probability (as in the Full and End-of-trial conditions).

In sum, two similar social comparison processes—both directed at achieving a favorable comparison of abilities—are consistent with the

full pattern of our present results, including the attenuated motivation gain observed in the Delayed-feedback condition. The latter attenuation effect could also be attributed to (a) lower concern about being revealed as the ‘weak link’ in the team (b) a risk of wasted effort or (c) reduced risks of an unfavorable social evaluation in this condition. It will require additional research to tease apart whether and to what degree each of these explanations contributes to the attenuation effect observed here.

Impression management concerns and the Köhler effect Another notion that is quite consistent with the present results is that the Köhler effect may ultimately stem from a desire to create a favorable (or avoid creating an unfavorable) impression among others. Each decrease in motivation we obtained was accompanied by a decrease in the availability of performance-relevant information to one’s partner. In the Full and End-of-trial feedback conditions, the participant’s status as the less capable group member was immediately evident as soon as he quit. In the Delayed-feedback condition, with its attenuated motivation gain, others could only discover who failed first (and even that revelation would only be made sometime after the trial was completed) but not when he failed. And in the No-feedback condition, which produced no motivation gain, no one could tell who quit first or how long he worked.

To further explore the role of impression management concerns, it would be informative to systematically vary ‘who knows what when’ across all possible sources of evaluation—the self, fellow group members, and non-performing observers (e.g. the experimenter), much as Harkins and Szymanski did in their productive program of research on the role of evaluation for social loafing (e.g. Harkins & Szymanski, 1989; Szymanski & Harkins, 1993). This would also include asymmetric feedback conditions in which different evaluators had access to different information; for example, a condition in which only P has full performance feedback of dyad members, but his/her partner and the experimenter do not.

Note that this emphasis on impression management need not contradict or compete with the other explanations of the Köhler effect. Rather, it may complement those explanations. For example, concern with the indispensability of one’s contribution to the group may reflect concern with how others would evaluate an indispensable group member who performed poorly. And one may try to keep up with a more capable teammate precisely because one is concerned about how one would be evaluated if one could not keep up. If so, then one insight of our study may be that group members are most likely to exert supranormal levels of effort on their group’s behalf when there are salient *social* incentives for doing so. If so, maximizing effort in work groups, teams, and organizations may depend as much or more on social psychological factors (such as establishing a favorable impression in others) as on more traditional, non-psychological factors (e.g. task engagement; pay).

Conclusions and qualifications

In this study, we systematically varied what group members knew about one another’s ability and when they knew it as they worked at a conjunctive group task. Besides establishing some informative boundary conditions on the Köhler motivation gain effect, our results show that certain explanations are insufficient to fully account for the effect. Specifically, the effect cannot be attributed solely to (a) conjunctive task demands per se making one feel indispensable to the group, or (b) using a more capable member’s performance to set a high performance goal. Our results also increase the plausibility that certain social comparison processes contribute to the effect. Specifically, our results are consistent with two such processes: (a) attempts to prevail in an intragroup competition with one’s partner, and (b) attempts to see just what one is capable of (to evaluate one’s ability) and to meet (or surpass) available social comparison targets (if possible, to show that one has relatively high ability). In addition, results from other studies (Hertel et al., 2000, Exp. 2; Kerr et al., 2005) as well as certain

results from the present study, also suggest that the effect is stronger when one's contribution is clearly indispensable to the group, even when opportunities for social comparison are held constant.

A running controversy in the group motivation area asks, 'what is the most appropriate baseline against which to contrast group-member motivation?' Some (e.g. Karau & Williams, 1993; Williams & Karau, 1991) have suggested that coactors—individuals working independently but in the 'presence' of one another, in one or another sense (e.g. visible to one another, able to monitor one another's performance)—provide the most appropriate baseline. Others (e.g. Markus, 2001; Todd, Kerr, Messé, & Seok, *in press*) have argued that isolated individuals represent better baselines, because coaction conditions can include some of the psychological processes available in groups (e.g. the opportunity for social comparison). Clearly, in this paper we have used individuals as the baseline to define our motivation gain effects. However, in other, in progress work (Kerr et al., 2005) we have run coactor conditions in which the participant worked (and was rewarded) individually, but got the same feedback about another worker as a participant in the present conjunctive dyads (i.e. the coactor always outperformed the participant). This coactor condition tends to produce a reliable but weaker motivation gain than conjunctive/dyad conditions. The clear implication is that the Köhler effect is partially driven by social comparison (present in both coaction and conjunctive conditions) and partially by an indispensability mechanism (additionally present in the conjunctive/dyad condition). The picture that is beginning to emerge is that there is probably not a single, unitary explanation for the motivation gain phenomenon first discovered by O. Köhler. A number of psychological processes are likely to be involved (e.g. social comparison, perceived indispensability of effort, perceived efficiency of effort, striving to create a favorable impression in others).

It also seems likely that different group or task contexts will alter the relative importance of these processes. For example, it seems likely that attempts to achieve a favorable ability

comparison would be stronger in contexts where one's ego is threatened (e.g. a male being outperformed by a female at a task requiring physical strength; Lount et al., 2000). It seems likely that attempts to fulfill one's role as the group's indispensable member would be stronger in contexts where one cared more about the group, one's standing in the group, and its success (e.g. in highly cohesive groups; when group identification was strong; Halsam, 2001, Chapter 9). The immediate challenge is to better understand exactly when and how these processes operate. Through such research, we may attain a more comprehensive and integrated understanding of how and why group work can be both motivating and demotivating.

Notes

1. This motor persistence task is similar in many respects to the task originally used by Köhler (1926, 1927), and has a number of important features (e.g. performance depends on little beside effort; performance can safely be assumed to be related monotonically to effort; cf. Hertel et al., 2000). However, none of the currently viable explanations for the Köhler motivation gain effect suggest that it is restricted to such simple motor tasks. And indeed, the effect has been replicated with a reasonably complex cognitive task (Hertel et al., 2003).
2. These instructions ensured that the extrinsic incentive was the same for participants in all conditions, under plausible assumptions (viz. that the coacting participant would score comparably to the participant; that both dyad members would receive the same team score).
3. Thanks to Garold Stasser for suggesting this interesting possibility.

Acknowledgments

This research was supported in part by an NSF grant (Motivation Gains in Performing Groups; BCS-9974664) to the first two authors. Portions of the reported research were presented at the Midwestern Psychological Association Annual Meeting, May, 2002. We wish to thank Dong-Heon Seok, Brendan Baird, and Robert Lount for their help conducting this research. Thanks also go to Garold Stasser and to two anonymous reviewers of an earlier draft of this paper for their helpful suggestions.

References

- Allport, F. H. (1962). A structuronomic conception of behavior: Individual and collective: I. Structural theory and the master problem of social psychology. *Journal of Abnormal and Social Psychology*, 64, 3–30.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7, 117–140.
- Guzzo, R. A., & Dickson, M. W. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology*, 47, 307–338.
- Hackman, J. R. (2002). *Leading teams: Setting the stage for great performances*. Cambridge, MA: Harvard Business School Press.
- Halsam, S. A. (2001). *Psychology in organizations: The social identity approach*. London: Sage.
- Harkins, S. G. (2001). Social influence effects on task performance: The ascendancy of social evaluation over self-evaluation. In K. D. Williams and J. Forgas (Eds.), *Social influence: Direct and indirect processes* (pp. 271–292). New York, NY, US: Psychology Press.
- Harkins, S. G., & Petty, R. E. (1982). Effects of task difficulty and task uniqueness on social loafing. *Journal of Personality and Social Psychology*, 43, 1214–1229.
- Harkins, S. G., & Szymanski, K. (1989). Social loafing and group evaluation. *Journal of Personality and Social Psychology*, 56, 934–941.
- Hertel, G., Deter, C., & Konradt, U. (2003). Motivation gains in computer-mediated work groups. *Journal of Applied Social Psychology*, 33, 2080–2105.
- Hertel, G., Kerr, N. L., & Messé, L. A. (2000). Motivation gains in performance groups: Paradigmatic and theoretical developments on the Köhler effect. *Journal of Personality and Social Psychology*, 79, 580–601.
- Karau, S. J., & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, 65, 681–706.
- Kelley, Harold H., Holmes, John G., Kerr, Norbert L., Reis, Harry T., Rusbult, Caryl E., & Van Lange, Paul A. M. (2003). *An atlas of interpersonal situations*. New York: Cambridge University Press.
- Kerr, N. L. (1990, October). *Reflections on group productivity*. Discussant in the Symposium 'Group Performance' (W. Stroebe, Chair), at the Annual Convention of the Society of Experimental Social Psychology, Buffalo, NY.
- Kerr, N. L., & Bruun, S. (1983). The dispensability of member effort and group motivation losses: Free rider effects. *Journal of Personality and Social Psychology*, 44, 78–94.
- Kerr, N. L., Messé, L. A., Seok, D. H., Sambolec, E., & Park, E. S. (2005). *Psychological mechanisms underlying the Köhler motivation gain effect*. Unpublished manuscript, Michigan State University.
- Köhler, O. (1926). Kraftleistungen bei Einzel- und Gruppenarbeit [Physical performance in individual and group situations]. *Industrielle Psychotechnik*, 3, 274–282.
- Köhler, O. (1927). Über den Gruppenwirkungsgrad der menschlichen Körperarbeit und die Bedingung optimaler Kollektivkraftreaktion [On group efficiency of physical labor and the conditions of optimal collective performance]. *Industrielle Psychotechnik*, 4, 209–226.
- Lount, R. B., Jr., Messé, L. A., & Kerr, N. L. (2000). Trying harder for different reasons: Conjunctivity and sex differences as bases for motivation gains in performing groups. *Zeitschrift für Sozialpsychologie*, 31, 221–230.
- Markus, M. J. (2001, May). *Rethinking social loafing: A task motivation hypothesis*. Paper presented at the Annual Meeting of the Midwestern Psychological Association, Chicago, IL.
- Messé, Lawrence A., Hertel, Guido; Kerr, Norbert L., Lount, Robert B.Jr, & Park, Ernest S. (2002). Knowledge of partner's ability as a moderator of group motivation gains: An exploration of the Koehler discrepancy effect. *Journal of Personality and Social Psychology*, 82(6), 935–946.
- Platt, J. R. (1964). Strong inference. *Science*, 146, 347–353.
- Shepperd, J. A. (1993). Productivity loss in performance groups: A motivation analysis. *Psychological Bulletin*, 113, 67–81.
- Steiner, I. D. (1972). *Group process and productivity*. New York: Academic Press.
- Stroebe, W., Diehl, M., & Abakoumkin, G. (1996). Social compensation and the Köhler effect: Toward a theoretical explanation of motivation gains in group productivity. In E. Witte & J. Davis (Eds.), *Understanding group behavior: Consensual action by small groups* (Vol. 2, pp 37–65). Mahwah, NJ: Erlbaum.
- Szymanski, K., & Harkins, S. G. (1993). The effect of experimenter evaluation on self-evaluation within the social loafing paradigm. *Journal of Experimental Social Psychology*, 29, 268–286.
- Todd, A. R., Kerr, N. L., Messé, L. A., & Seok, D.-H. (in press). *Social compensation: Fact or social comparison artifact? Group Processes & Intergroup Behaviour*.

- Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.
- Weldon, E., & Gargano, G. M. (1988). Cognitive loafing: The effects of accountability and shared responsibility on cognitive effort. *Personality and Social Psychology Bulletin*, *14*, 159–171.
- Williams, K. D., & Karau, S. J. (1991). Social loafing and social compensation: The effects of expectations of co-worker performance. *Journal of Personality and Social Psychology*, *61*, 570–581.
- Zajonc, R. B. (1965). Social facilitation. *Science*, *149*, 269–274.

Paper received 25 March 2004; revised version accepted 2 November 2004.

Biographical notes

- NORBERT L. KERR is professor of psychology at Michigan State University. His research interests focus on small group performance and decision making, cooperation in social dilemmas, and psychology and the law.
- LAWRENCE A. MESSÉ was professor of psychology at Michigan State University. His research interests included the personal bases of interpersonal behavior, moderators of interpersonal conflict, group performance, and minority influence. He died on December 1 2004.
- ERNEST S. PARK is a postdoctoral research scholar at North Dakota State University. His primary interests are in group performance and decision making, affect in groups, and coordination in teams.
- ERIC SAMBOLEC is a doctoral student in social psychology at Michigan State University. His research interests include group and individual motivation, performance attributions, and sport psychology.