Journal of Sport & Exercise Psychology, 2011, 33, 506-526

Buddy Up: The Köhler Effect Applied to Health Games

Deborah L. Feltz, Norbert L. Kerr, and Brandon C. Irwin
Michigan State University

The present investigation examined the Köhler motivation gain effect in a health game using an absent partner, presented virtually. The Köhler effect occurs when an inferior team member performs a difficult task better in a team or coaction situation than one would expect from knowledge of his or her individual performance. The effect has been strongest in conjunctive task conditions in which the group's potential productivity is equal to the productivity of its least capable member. Participants were randomly assigned to one of four conditions (individual control, coaction, additive, and conjunctive) in a 4 (conditions) × 2 (gender) factorial design and performed a series of isometric plank exercises within an exercise game. They performed the first series of five exercises alone holding each position for as long as they could, and, after a rest period, those in the partner conditions were told they would do remaining trials with a same-sex virtual partner whom they could observe during their performance. The partner's performance was manipulated to be always superior to the participant's. Results showed that task persistence was significantly greater in all experimental conditions than in the individual control condition. The conjunctive condition was no more motivating than either the additive or coactive conditions. Results suggest that working out with virtually present, superior partners can improve persistence motivation on exercise game tasks.

Keywords: conjunctive task, dyad exercise, exergame, persistence motivation, virtual partner

The present research was designed to determine whether recently documented motivation gains in task groups (dyads in particular) could be harnessed to improve people's motivation in health video games using a virtually presented partner. If people's motivation can be improved under such conditions, increasing the intensity and duration of exercise by participating with such a partner, they will realize better health outcomes than if they exercise alone.

There are a number of social and psychological factors that influence motivation to exercise (Franzini et al., 2009; USDHHS, 2008). These include social support from health professionals, family, and friends (Coleman, Cox, & Roker,

---

Deborah L. Feltz is with the Department of Kinesiology, Norbert L. Kerr is with the Department of Psychology, and Brandon Irwin is with the Department of Kinesiology, Michigan State University, East Lansing, MI.

2008; Zakarian, Hovell, Hofstetter, Sallis, & Keating, 1994), social modeling of physical activity (Feltz & Riessinger, 1990; Fox, Rejeski, & Gauvin, 2000), coexercisers (Carron, Hausenblas, & Mack, 1996), group exercise programs (Dishman & Buckworth, 1996), and access to and convenience of exercise facilities (Sallis et al., 1990). In addition, enjoyment of physical activity and self-efficacy for overcoming exercise barriers have consistently been linked to exercise adherence (McArthur & Raedeke, 2009; McAuley, 1993; Sallis, Prochaska, Taylor, Hill, & Geraci, 1999).

Research also suggests that some social environments are more appropriate for fostering motivation and quality exercise experiences. For example, researchers have consistently found group exercise leads to higher exercise adherence than individual exercise programs (Dishman & Buckworth, 1996). Specifically, group exercise programs are related to higher enjoyment and levels of social support as well as increased intention to continue exercising. On the other hand, structured group exercise programs present special problems for those with social physique anxiety (Bain, Wilson, & Chaikind, 1989) or those who lack the time and/or resources to join an exercise group. Moreover, previous models of group exercise rarely if ever introduce any real interdependence between exercisers (e.g., create teams whose progress and or outcomes are mutually determined).

However, there is now considerable basic research on motivation gains in task groups showing that under the right conditions, effort at demanding physical tasks (e.g., exercise) can be boosted via a number of social psychological processes that focus on social comparison, competition, and obligation. One well-studied motivation gain phenomenon that has promise for motivating greater effort in exercise is the *Köhler effect* (Hertel, Kerr & Messé, 2000; Köhler, 1927; Messé, Hertel, Kerr, Lount, & Park, 2002). The Köhler effect was first described in the 1920s by the German industrial psychologist Otto Köhler. Studying male members of a Berlin rowing club, Köhler found that dyads could perform a taxing physical task (viz., doing as many standing bicep curls as possible) longer than one would expect from knowledge of the dyad members' performances at a comparably difficult individual version of the task. The demands of Köhler's dyad task meant that the group could persist no longer than its weaker member; once that weaker member was exhausted and quit, it was impossible for the stronger member to continue.

Such group tasks—in which the group's potential productivity is equal to the productivity of its least capable member—are commonly referred to as *conjunctive* tasks (Steiner, 1972). So, essentially Köhler demonstrated that the weaker member of a dyad will push him- or herself harder—beyond their usual performance limits—when paired with someone stronger in a conjunctive persistence task. Köhler also found that this motivation gain was moderated by the discrepancy between dyad partners' abilities—the motivation gain was largest when this discrepancy was moderate (in Köhler's studies, the maximum gain occurred when one partner was able to persist individually about 1.4 times longer than the other).

Recent research (e.g., Kerr et al., 2007) reveals that there are at least two mechanisms underlying the Köhler effect. The first stresses social comparison processes (Festinger, 1954; Seta, 1982; Stroebe, Diehl, & Abakoumkin, 1996; Suls & Wheeler, 2000). When confronted with a more capable partner on an ambiguous but valued task, the weaker partner may revise his or her personal performance goal upward. A variant on this notion suggests that doing as well or better than the partner—successful competition—becomes a salient goal. Although there are interesting differences between the former, goal-setting version and the latter, intragroup-competition version, both versions hold that the opportunity for comparison of performance levels is crucial to produce a gain in motivation. Such opportunities can arise even when performers are not actually working together as a group (e.g., when they are coacting, that is, when two people exercise in one another's presence). Hence, this explanation holds that working together in an interdependent group is not essential for a Köhler motivation gain.

The second mechanism stresses the indispensability of one's efforts for one's group. As Instrumentality × Value models of motivation suggest (e.g., Vroom, 1964; Karau & Williams, 1993), task motivation is likely to be enhanced when one sees one's efforts as being highly instrumental in achieving highly valued outcomes. Under conjunctive task conditions, the group's performance (and other performance-mediated outcomes, such as one's reputation in the group) is highly contingent on the weaker member's effort; that is, the weaker member's efforts are indispensable for group success. Note that such contingencies depend largely on the demands of the task—it is the conjunctive nature of the group task that makes the weaker member's efforts particularly indispensable.

The Köhler effect was first observed in the training room of amateur athletes (rowers). More recently, it has been studied in ad hoc laboratory groups for a variety of motor and cognitive tasks. A series of experiments (see Kerr et al., 2007), as well as a meta-analysis (Weber & Hertel, 2007), have shown that the gain attributable to the indispensability mechanism—estimable from the difference in motivation observed under coactive vs. conjunctive task conditions—tends to be relatively larger for females than for males. At least in the laboratory, it appears that successful social comparison or competition is a relatively more important contributor to the Köhler effect for males, whereas concerns about not letting one's partner or the group down is relatively more important among females than males.

Although working coactively or conjunctively with a partner in ad hoc laboratory groups has shown motivational gains for the least capable member of the team, there are some practical obstacles to implementing the principles in exercise settings. For instance, trying to find the optimally matched partner can be difficult, individuals can become discouraged if they believe they can never keep up with their partner, or on the other hand, become bored if their partner is always slower. However, the use of a virtually present partner has the benefit of being more practical than trying to find the best matching live partner, of being helpful to those with social physique anxiety, and in being easier to adapt to changing performance improvements of the matched partner.

None of the extant health games (e.g., Wii Fit, XaviX J-Pad, PS-2's EyeToy: Kinetic) use virtual partners that take much advantage of the potential of group dynamics to motivate participation or incorporate the critical design features suggested by previous research and theory on the Köhler effect (viz., immediate feedback on performance of one or more other players, the ability to control the discrepancy in abilities of players, and the indispensability of individual player effort for determining group outcomes). And, although health video games have shown some health benefit in terms of increased caloric expenditure and cardiorespiratory endurance (e.g., Porcari, Schmidt, & Foster, 2008), the interpersonal and social concerns (for comparing favorably with others, or for not letting a partner or teammates down) that can be incorporated into these games have the potential to add equally powerful new sources of motivation.

Thus, our investigation sought to determine how the intensity and duration of exercise during participation in a health game with a virtual partner can be enhanced by harnessing social psychological mechanisms discovered in group dynamics research (e.g., Kerr et al., 2007; Köhler, 1927; Weber & Hertel, 2007). Participating in exercises in this collaborative way could also enhance participants' enjoyment and interest in continuing the game, self-efficacy, and intention to exercise in the future. Because these three variables have been associated with exercise adherence (e.g., McArthur & Raedeke, 2009; McAuley, 1993), we were also interested in the game's effect on them.

Based solely upon the overall patterns observed in previous laboratory research employing effort-sensitive physical and cognitive tasks, we advance the following tentative hypotheses:

- H1a: The duration of exercise will be greater with a moderately more capable coacting partner than when exercising alone (due to the social comparison mechanism), and greater still with a moderately more capable teammate exercising under conjunctive team task demands (due to the indispensability mechanism). That is, individuals < coaction condition < conjunctive condition.
- H2a: The motivation gain observed (i.e., duration of exercise) under coaction conditions in health games will be larger for males than for females, coactive-$\text{individuals}_{\text{males}}$ > coactive-$\text{individuals}_{\text{females}}$, whereas the additional motivation gain observed under conjunctive conditions will be larger for females than for males, conjunctive-$\text{coactive}_{\text{males}}$ < conjunctive-$\text{coactive}_{\text{females}}$.

However, aspects of the group exercise setting might qualify these hypotheses. For example, it has been shown (Kerr et al., 2007, Exp. 3) that priming competitiveness leads females to act more like males, that is, to show less of an indispensability component for the Köhler effect. Video games, including exercise games, often include playing with a partner, and the partners compete with each other. To the extent that exercising with others in a video-game context strongly primes competition, we might observe similar effects in our health-game context. This would result in a different, alternative pattern of results, as follows:

- H1b: The duration of exercise will be greater with a moderately more capable coacting partner than when exercising alone (due to the social comparison mechanism), but no greater with a moderately more capable teammate exercising under conjunctive team task demands. That is, individuals < coaction condition = conjunctive condition.
- H2b: Males and females will show the same pattern of motivation gain effects under coactive and conjunctive task conditions.

Finally, results are somewhat mixed about how hard group members work when the task demands are additive (i.e., the group score is just the sum of individual member contributions). Some studies (e.g., Hertel, Deter, & Konradt, 2003; Hertel, Kerr, Scheffler, et al., 2000, Exp. 2; Hertel, Kerr, & Messé, 2000, Exp. 2) find no significant motivation gain under additive conditions, even when social comparison between group members is possible. Weber and Hertel's (2007) meta-analysis, by contrast, finds a comparable motivation gain under additive and coactive work conditions. Given this ambiguity and the fact that ours is the first study looking at the Köhler effect in exercise groups, we approached the contrast of additive and coactive work conditions as an interesting but open research question.

Method

Participants

Although the population of people who might benefit from this motivation-gain approach to health games is potentially quite large—including children and younger and older adults who want or need to improve their fitness and health by playing health games—in this initial phase of our research, we focused our attention on healthy young adults (college students). Thus, participants were 181 college students (M_{age} = 20.10, *SD* = 1.75) recruited from introductory psychology and kinesiology courses at a large Midwestern university. Students were given course credit for their participation. Both male and female experimenters conducted sessions throughout the experiment.

Design

The experiment used a 4 (work condition: individual control, coactive, additive, conjunctive) × 2 (gender) × 2 (block: first vs. second) design with repeated measures on the last factor. Within each trial block, participants performed five isometric exercises: front plank, side plank (left), one leg plank (left), side plank (right), one leg plank (right) (see below for details). One participant took part in each session that lasted about 1 hr. Within each gender, participants were randomly assigned into one of the four work conditions (with a proviso that fewer participants were run in the exploratory additive condition, to concentrate statistical power in those conditions of primary interest). As a result, there were 49 in the individual condition (24 males, 25 females), 56 in the coactive condition (25 males, 31 females), 28 in the additive condition (13 males, 15 females), and 48 in the conjunctive condition (25 males, 23 females). The primary dependent variable was persistence in the exercises, indexed by the total time the exercises were held across the trial.

Health Game and Task

The health game used for this study was an exercise video game (exergame) designed for the Playstation 2 (PS2) gaming module. The software used was EyeToy: Kinetic, a game that offers a variety of fitness activities (e.g., yoga, strengthening exercises, combat exercises). This particular software operates in conjunction with an additional accessory called the Eye Toy, designed specifically for the PS2 system. The Eye Toy is essentially a small camera that connects to the PS2 system via a USB cable and allows images of the user to be displayed on the TV monitor and interact with virtual environments supported by the software.

The abdominal plank exercises within the strength training module of the EyeToy: Kinetic software were used for this experiment. These are a type of bodyweight exercise where participants are required to suspend their own body weight using their abdominal muscles. These exercises are also isometric in nature and require very little coordination, and thus are highly effort based. Each exercise targeted the abdominal muscles, but there were slight differences between each. On the first exercise, participants were face down on a cushioned mat, with legs extended straight, and lifted their body upward by resting their elbows and toes on the mat and using their abdominals to lift their body. In this way, the body was in a straight line, the spine was directly in line with their head and legs and nothing was touching the ground except for the elbows, forearms, and toes. In a similar fashion, the second exercise achieved the same elevated position, but the participant was on the left side with only the left forearm and left foot on the ground, emphasizing the use of the outer abdominal muscles. The third exercise was the same as the first exercise except that the participant had the left leg raised in the air and thus was balancing on only the right foot, which emphasized the lower abdominal region. The fourth exercise was the same as the second, except the participant performed this on the right side. The fifth exercise was the same as the third, except the participant performed this with the right foot in the air (see Figure 1). Participants performed each exercise once within each of two blocks.

Measures

Performance. Performance was the total number of seconds that the exercise was held. Block scores were calculated by taking the summed total of the five exercises within each trial, justified by a high intraclass correlation coefficient for each trial (Trial 1 = .83, Trial 2 = .85).

Heart Rate. Although the plank exercises were isometric in nature, the heart rate (HR) should still increase the longer one persists at the task. We used HR to provide a physiological measure of participants' levels of exertion across conditions. The HR measure also allowed a comparison with subjective ratings of perceived exertion (RPE) for participants across conditions. HR was measured using a Polar E600 heart rate monitor. The monitor operates by recording electrical signals produced by the heart with each beat. In this manner, HR was sampled and summarized in 5-s epochs. Heart rate was measured during exercise (i.e., not during rest periods between exercises and between blocks).[1] Overall HR scores for each trial were calculated by averaging HR scores from each exercise within each trial.

Self-Efficacy. Task self-efficacy (SE) was measured with a scale developed specifically for this study. The measures contained five items, each corresponding to one of the five exercises within each trial. All items were preceded by the stem "What is the number of seconds that you are completely confident you can hold:" followed by "The first exercise," "the second exercise" and so on for each of the five exercises. Respondents wrote in the number of seconds in a blank box following each item. The questionnaire was administered at three time points. Once before Trial 1 (after the participant had watched a brief instructional video demonstrating the exercises), a second time after performing the exercises for Trial 1, and again after the second trial of exercises. A total SE score for each trial was calculated by taking the sum of the five items within each trial.

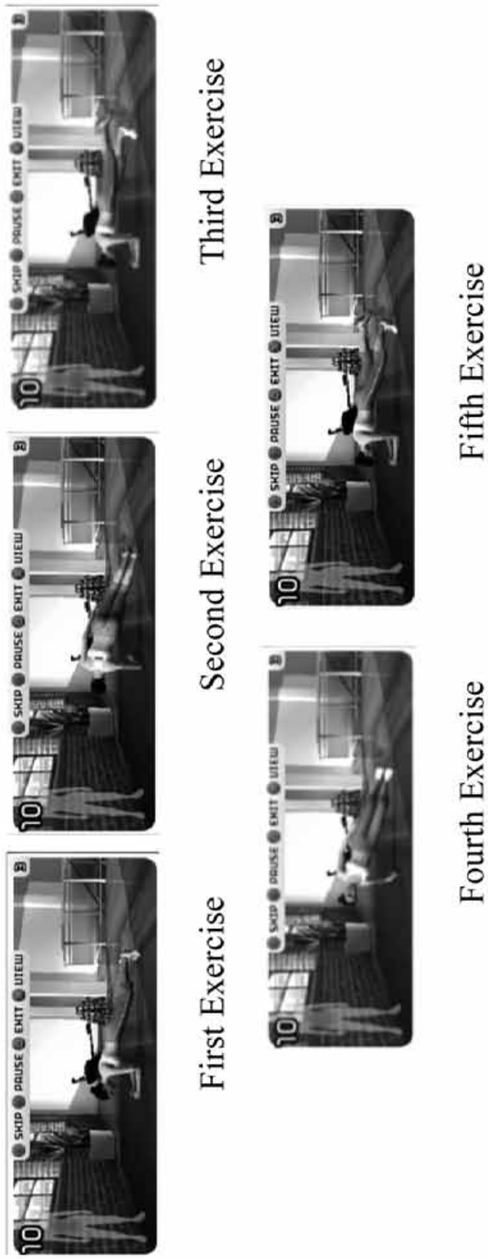

Figure 1 — Images of exercises performed.

Ratings of Perceived Exertion. Perceived exertion was measured using the 6–20 version of the Borg (1998) RPE scale. The scale ranges from 6 to 20, where 6 means *no exertion at all* and 20 means *maximal exertion*. Participants were asked to rate their exertion at the end of each exercise, with particular reference to their perceived exertion at the moment right before the end of the exercise.

Enjoyment. Enjoyment was measured at the end of Trial 2 using a short eight-item version of the Physical Activity Enjoyment Scale (PAES; Kendzierski & DeCarlo, 1991). Each item was rated on a 7-point bipolar scale beginning with the stem "Please rate how you feel at the moment about the physical activity you have been doing according to the following scales" (e.g., 1 = *I loved it*; 7 = *I hated it*). Previous studies have shown high correlations with the complete scale (r = .94; Raedeke, 2007) and strong reliability (α = .91; McArthur & Raedeke, 2009).

Intention to Exercise in the Future. Intention to exercise was measured at the end of Trial 2 with one item asking participants to respond to the following statement: "I intend to exercise tomorrow for at least 30 minutes" on a scale of –3 (*Not at all true for me*) to +3 (*Completely true for me*). Previous researchers have demonstrated the validity of a single item in measuring this construct within the exercise domain (Chatzisarantis, Hagger, Biddle & Smith, 2005; Rhodes & Courneya, 2005).

Postexperimental Questionnaire. Besides some questions checking participants' understanding of the instructions and procedures, there were questions probing their interest in participating in a future exercise study like the present one, a rating of task difficulty, and a rating of effort expended on the task, each made on 8-point scales.

Procedures

Before conducting this study, permission was obtained from the institutional review board. Before each session, an experimenter made sure that none of the participants had any disabling injuries to their arms, shoulders, back, or legs. Once an informed consent form was signed, participants were asked to remove any wrist jewelry/watches. The experimenter then demonstrated how to put on the HR monitor. Participants first watched a brief instructional video including images from the PS2–EyeToy: Kinetic software in which a virtual trainer demonstrated the five exercises. A baseline measure of self-efficacy was then taken. An incentive for good performance was then described by the experimenter, as has been typical of tests of the Köhler effect (e.g., Kerr et al., 2007). In the current study, persistence scores from all trials translated into lottery tickets (one ticket per second persisted; the lottery winner would receive a membership at the university's fitness center, worth approx. $80.00). Participants were not told the exact number of trials they would be performing, but only that they would perform a series of trials. All participants then performed the first block of exercises, holding each exercise as long as they could and with 30-s rest periods between each exercise. A button on the wrist monitor was pressed by the participant at the start of each exercise and again when the participant quit.[2] Immediately after each exercise, the participant announced his or her perceived exertion on the 15-point Borg scale. All participants were given veridical feedback from the experimenter on their performance (i.e., the average number of seconds they held each exercise).

The work condition manipulations were introduced at this point. Participants in the individual control condition simply rested for 10 min. Participants in the remaining experimental conditions were told that another participant was being tested simultaneously at another laboratory, and that the two participants would be able to see one another over a "Skype-like" video connection during future trials. The participants then met briefly with that other, same-sex participant in a controlled Skype-like interaction (we will refer to that other participant hereafter as *the partner*). In reality, the partner was an experimental confederate whose side of the interaction was prerecorded. The interaction lasted for approximately 2 min. The partner first shared his or her name, year in school, career plans, and favorite television show. After the confederate was finished speaking, the participant shared the same information as the confederate "listened" on the other end (partners were trained to have a very bland countenance throughout the recording, as to avoid any grand facial expressions that might influence the subject's attitude toward the partner). After the interaction, participants were also given bogus feedback on how well the partner had done on the first trial. That feedback was 1.4 times the participant's own actual performance; previous work (e.g., Köhler, 1927; Messé et al., 2002) indicates that this approximate level of partner ability superiority is optimal for producing the Köhler motivation gain effect.

The work conditions differed in the interdependence between the participant and the partner. In the coaction condition, participants were told that they could watch one another over a video link as they performed each exercise. However, each of their chances in the upcoming lottery was solely dependent on their individual performances—there was no task interdependence. In the remaining two conditions (additive and conjunctive), participants were told that they would be an exercise team and both teammates would earn the same number of lottery tickets, to be determined by the team's score. In the additive condition, the team score would simply be the average of the two teammates' individual persistence scores. In the conjunctive condition, the team score would be the persistence score of the first teammate to quit an exercise. Following the manipulation, participants responded to a self-efficacy questionnaire.

Block 2 then commenced. In the individual control condition, the participant could only see him- or herself on the screen, as during Block 1. In the remaining conditions, the participant could see the partner's image (which was actually prerecorded) before and during the exercise in addition to his or her own; the participant believed that the partner could likewise see his or her (the participant's) image. The images available to the participant suggested that she or he was always the first to quit each exercise. The video link was allegedly frozen as soon as either teammate quit an exercise and until just before the start of the next exercise; hence, the participant only knew that the partner was able to persist longer, but not how much longer. In actuality, this period lasted 30 s, the same amount of rest time allotted between exercises during Block 1.[3]

After Block 2 was over, the participant completed a series of questionnaires (self-efficacy, intention to exercise, enjoyment of physical activity, and manipulation checks). The participant was then debriefed, thanked, and excused.

Results

Preliminary Analyses

Our initial analyses focused on the individual controls. Any performance drop in this condition between Block 1 and Block 2 estimates the effect of fatigue and boredom with the task, and provides an important baseline against which we can compare the remaining experimental conditions. We also wanted to see if the sex of the experimenter was more than a nuisance variable.

To provide an overall estimate of persistence, the number of seconds each participant held his or her plank exercise was summed across the five discrete exercises to produce two persistence scores, one at Block 1 and a second at Block 2. These scores were entered into a 2 (block) × 2 (sex of participant) × 2 (sex of experimenter) mixed ANOVA with repeated measures on the first factor. Only two effects were significant. Males persisted longer *(M* = 294.92 s, *SD* = 96.17) than females (*M* = 201.038 s, *SD* = 95.01) at these tasks, $F(1,45) = 7.28$, $p = .011$, $\eta_p^2 = .139$. More interestingly, all participants persisted longer at Block 1 *(M* = 275.90 s, *SD* = 117.01) than at Block 2 (*M* = 220.04 s, *SD* = 98.09), $F(1,45) = 47.75$, $p < .001$, $\eta_p^2 = .515$). This 55.86-s fatigue effect was not moderated by participant gender (interaction $F < 1$) and the gender of the experimenter showed no significant effects (and was hence ignored in subsequent analyses).

Main Motivation Gain Analyses

Exercise Persistence. There are, of course, individual differences in fitness and strength that we wish to control for. In some early studies (e.g., Hertel, Kerr, & Messé, 2000; Kerr et al., 2007), Block 1 performance was used as a baseline and fatigue-corrected difference scores (i.e., Block 2 – Block 1 + fatigue correction) were the primary dependant variable. An alternative that is less vulnerable to certain problems that arise from the use of difference scores (e.g., Edwards & Parry, 1993) is to use Block 1 scores as a covariate in the analysis of Block 2 scores. This method simultaneously controls for individual differences in strength and the overall mean fatigue effect. Here, we present the results using the latter method (although both methods produce the same pattern of results).

Block 2 persistence scores were analyzed in a 2 (participant gender) × 4 (work condition: individual, coactive, additive, conjunctive) ANCOVA with Block 1 persistence scores as a covariate. First, and unsurprisingly, Block 1 scores predicted Block 2 scores, covariate $F(1,168) = 358.19$, $p < .001$, $\eta_p^2 = .681$. Second, and of primary interest was a significant work condition effect, $F(3,168) = 11.67$, $p < .001$, $\eta_p^2 =. 172$. The relevant work condition means and standard deviations are provided in Table 1 and are plotted in Figure 2. As the figure shows, compared with participants in the individual condition, participants in all three experimental conditions persisted longer. This is apparent from the confidence intervals in Figure 2, or from direct planned contrasts: individuals vs. conjunctive $t(168) = 4.20$, individuals vs. additive $t(168) = 3.37$, and individuals vs. coactive $t(168) = 5.67$, $p < .001$ for all contrasts. Although the mean corrected persistence score in the coactive condition (288.58 s) is slightly higher than the means in the additive (270.18 s) or conjunctive (273.11 s) conditions, these differences are not significant (via planned contrasts or inspection of confidence intervals).[4] So Hypothesis 1b is confirmed and Hypothesis 1a is disconfirmed in this group exercise context, and the equivalence of additive and coactive conditions suggested by Weber and Hertel's (2007) meta-analysis is corroborated. The overall work condition effect was not moderated by participant gender, $F(3,168) = 1.32$, *ns;* hence, Hypothesis 2b is confirmed and Hypothesis 2a disconfirmed. There was a weak trend for females to show greater Block 2 persistence ($M = 271.54$ s) than males ($M = 256.11$ s), but this difference was not significant ($p = .11$).

Heart Rate. Heart rate was used to provide a physiological measure of participants' levels of exertion across conditions. Parallel analyses of the average HR data were performed; that is, average HRs during Block 2 exercises were analyzed in a 2 (gender) × 4 (work condition: individual, coactive, additive, conjunctive) ANCOVA with Block 1 HRs as a covariate. The pattern of the data was identical to that obtained for the persistence data (see Table 1 and Figure 3). Block 1 scores predicted Block 2 scores, covariate $F(1,135) = 211.38$, $p < .001$, $\eta_p^2 = .610$, and there was a significant work condition effect, $F(3,135) = 4.25$, $p < .01$, $\eta_p^2 =. 086$. Again, the Block 2 HRs were significantly higher in the three experimental conditions than in the individual control conditions—individual vs. conjunctive $t(135) = 2.88$, $p < .01$; individual vs. additive $t(135) = 3.01$, $p < .01$; and individual vs. coactive $t(135) = 2.26$, $p < .025$—but not significantly different among the experimental conditions.

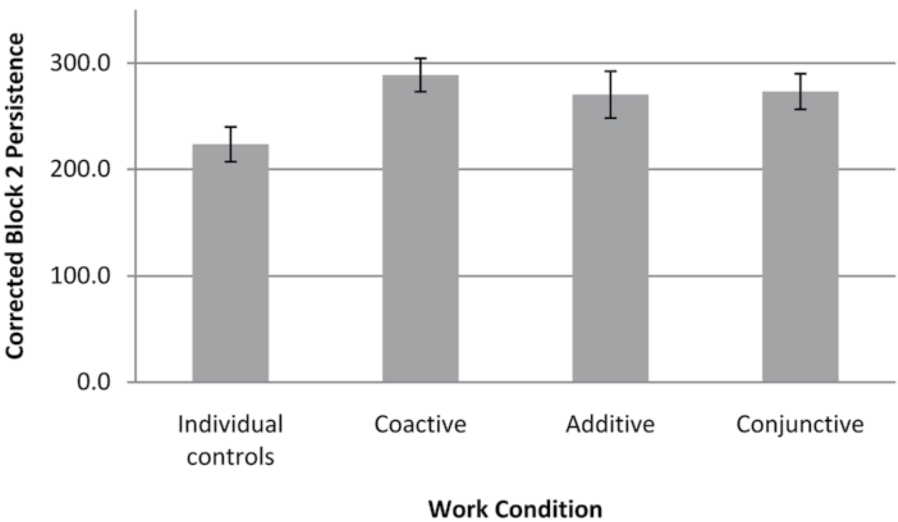

Figure 2 — Block 2 persistence (with Block 1 as a covariate).

Table 1 Condition Means and Standard Deviations

	Condition			
Variable	Individual Control	Coactive	Additive	Conjunctive
Block 2 persistence[a]	223.43 (57.96)	288.58 (58.49)	270.18 (57.99)	273.11 (58.06)
Block 2 heart rate[a]	114.10 (13.33)	120.53 (13.25)	124.74 (13.25)	122.73 (13.25)
Ratings of perceived exertion	14.54 (1.74)	13.98 (1.75)	14.42 (1.74)	14.56 (1.74)
Postexp. rating of effort expended	5.98 (1.23)	5.58 (1.32)	5.85 (1.38)	6.00 (1.26)
Self-Efficacy[a]	165.26 (107.17)	174.56 (107.85)	216.40 (107.20)	181.36 (107.45)
Enjoyment of task	3.88 (.92)	3.73 (.85)	3.65 (.75)	3.73 (.98)
Task difficulty	4.69 (1.69)	4.98 (1.26)	5.31 (1.23)	5.58 (1.26)
Intention to exercise	1.47 (1.52)	1.71 (1.61)	1.55 (1.78)	1.35 (1.85)
Willingness to participate again	5.87 (1.79)	5.30 (2.03)	5.31 (1.91)	5.07 (2.18)

Note. In repeated-measures ANOVAs and ANCOVAs, cell standard deviations estimated using within-Ss error term.

[a]Tabled means are estimated, adjusted for the effect of using Block 1 scores as a covariate.

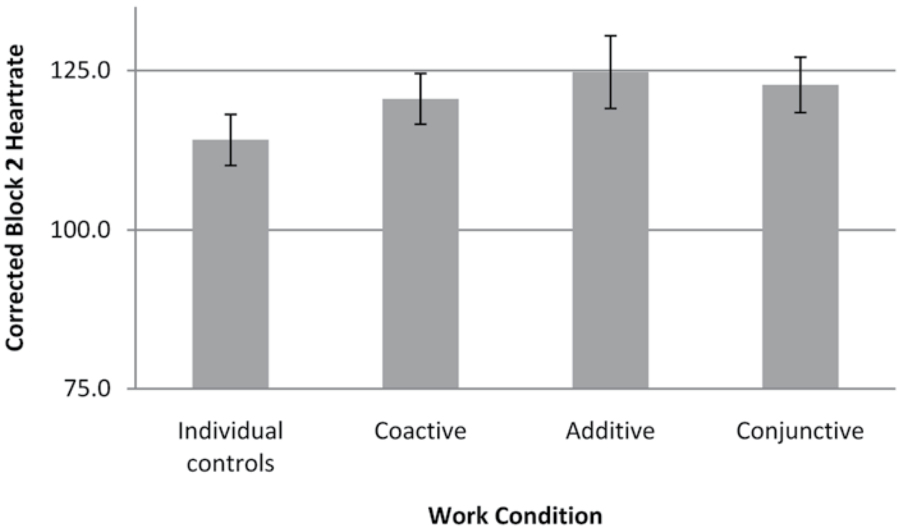

Figure 3 — Block 2 HR (with Block 1 as a covariate).

Ancillary Analyses

All variables collected at Block 1 and Block 2 (RPE and SE) were analyzed in a 2 (block) × 2 (gender) × 4 (condition) mixed ANOVA with repeated measures on the Block factor. Variables collected at the end of the experimental session (enjoyment, ratings of task difficulty, and intention to exercise in the future) were analyzed in 2 (gender) × 4 (condition) ANOVAs. Table 2 presents the correlations between persistence and self-report measures.

Ratings of Perceived Exertion. The only effects to emerge were a Block main effect, $F(1,170) = 39.48, p < .001, \eta_p^2 = .188$, in which participants reported greater exertion during Block 2 exercises (M = 14.62, SD = 1.67) than during Block 1 exercises (M = 14.13, SD = 1.69) (not surprising given their greater fatigue at Block 2), and a Gender main effect, $F(1,170) = 6.76, p = .01, \eta_p^2 = .038$, in which males reported greater exertion (M = 14.73, SD = 1.56) than females (M = 14.02, SD = 1.89).

Likewise, postexperimental ratings of how much effort had been expended doing the exercises showed no effect of work condition, $F(3,169) = 1.20$, *ns,* and a marginally significant effect of gender, $F(1,169) = 3.69, p = .057$, with males reporting somewhat greater effort (M = 6.04, SD = 1.19) than females (M = 5.66, SD = 1.37). Work condition showed no main or interaction effects. Thus, although HR and persistence data indicated that participants worked harder in the three partnered conditions, they did not perceive their work to be any harder.

Exercise SE. Self-Efficacy was assessed via the total number of seconds that participants estimated they could persist at the five plank exercises. As shown in previous research (Moritz, Feltz, Fahrbach, & Mack, 2000), initial (Block 1)

Table 2 Correlations Between Persistence and Self-Report Measures

	Block 1 Persistence	Block 2 Persistence	Enjoyment	Baseline SE	Post–Block 1 SE	Post–Block 2 SE	Block 1 RPE	Block 2 RPE	Intention
Block 1 persistence	—	.788**	.074	.304**	.533**	.471**	.092	.139*	.050
Block 2 persistence		—	.083	.198**	.475**	.529**	–.044	.034	.046
Enjoyment			—	.035	–.011	.009	–.019	.068	.337**
Baseline SE				—	.460**	.322**	.143*	.231**	–.002
Post–Block 1 SE					—	.784**	.074	.143*	–.075
Post–Block 2 SE						—	.028	.066	–.093
Block 1 RPE							—	.858**	–.068
Block 2 RPE								—	–.070
Intention to exercise									—

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

pre-performance SE scores were moderately associated with the actual performance during the first block ($r = .24$, $p < .001$); moreover, the mean estimate (227.38 s) was reasonably accurate (mean of actual Block 1 persistence = 268.75 s). The (Block 1) post-performance SE scores, informed by recent experience, were even better predictors of Block 2 performance ($r = .43$, $p < .001$). In addition, Block 1 performance significantly predicted post–Block 1 SE scores ($r = .50$, $p < .001$).

It was of interest to see if post-performance SE was affected by participants' work condition. This was examined in a 2 (block) × 2 (gender) × 4 (work condition) ANCOVA that used the pre–Block 1 SE score as a covariate. The only effect to emerge was a Block main effect, $F(1,171) = 8.74$, $p = .004$, $\eta_p^2 = .049$; participants were less sanguine about their prospects for persisting after Block 2 ($M = 166.27$ s, $SD = 113.79$) than after Block 1 ($M = 202.52$ s, $SD = 128.03$). But the work condition differences in performance were not paralleled by differences in SE for the exercise task.

Evaluation of the Exercise Task. An overall enjoyment measure was computed based on the seven items of the PAES scale. Neither gender nor work condition affected this judgment; the grand mean on the 7-point scale was 3.75 ($SD = .89$), significantly ($p < .05$) below the indifference midpoint, suggesting mild dislike of the exercise. Similarly, participants viewed the task as moderately difficult (grand mean = 5.14, $SD = 1.49$, significantly, $p < .05$, above the rating scale midpoint). There was a work condition effect as well, $F(3,169) = 3.30$, $p = .022$, $\eta_p^2 = .055$; the task was viewed as significantly more difficult (5.58) in the conjunctive condition than in the individual condition (4.69), with the remaining conditions falling between.

Intention to Exercise in the Future. Overall, at the end of the experimental session, participants expressed a positive intent to exercise for at least 30 min the following day. The grand mean was 1.52 ($SD = 1.68$) on the 7-point scale anchored by –3 (*Not at all true for me*) to +3 (*Completely true for me*); this was significantly ($p < .05$) above the midpoint of the scale. This intention was not affected by either gender or work condition. In a related vein, participants expressed a moderately high willingness to participate in another study like this one ($M = 5.39$, $SD = 2.00$, significantly above the midpoint of the 7-point scale), and again, to equal degrees across work conditions.

Discussion

The primary aim of our study was to examine the Köhler motivational effect using a virtually presented partner within a health video game context. Further, we were interested in comparing the mechanisms (social comparison and indispensability) under which the Kohler effect might improve exercise persistence. Our results showed that upward comparison with a more capable virtually presented partner (in all experimental conditions) was sufficient to produce a motivation gain. In fact, those who exercised with a more-capable partner persisted 53.86 s longer on average ($M = 277.29$ s) than the individual controls ($M = 223.42$ s). This represents a 24.1% increase—a considerable gain practically speaking for those who might be working on improving core stability in therapeutic or training settings. The fact that our effect was found with a virtually presented partner overcomes some of the practical obstacles of finding an optimally matched partner to exercise with at a particular location. For instance, prerecorded videotapes of different types of partners can be produced and manipulated to always be moderately more capable than the target exerciser. It should also be noted that previous research (Hertel, Niemeyer, & Clauss, 2008; Lount, Park, Kerr, Messé, & Seok, 2008) indicates that the physical presence of a coactor or teammate tends to boost both components of the Köhler effect. Thus, the effect we have observed with virtually present partners is likely to underestimate the magnitude of the effect that would be observed in face-to-face exercise teams. We must note, however, that in face-to-face exercise dyads, what is gained for the weaker could be lost for the stronger.

The motivation gains achieved with a more capable partner did not come at the expense of aversion to the task. Although all partnered participants had higher HRs and some (viz., in conjunctive condition) viewed the task as more difficult than individual controls, there was no evidence that they perceived they were working any harder (i.e., no differences in RPE), enjoyed the exercise less, or had lower SE at the task than controls. The lack of differences in RPE supports previous research showing that overweight individuals who control a video game while exercising expend significantly more energy without a related increase in RPE than those who exercise without the video game interaction (Fitzgerald et al., 2004; Haddock, Siegel, & Wikin, 2009). Researchers have suggested that being immersed in an enjoyable game/activity distracts one from the perceptions of greater exertion. We suggest that exercising with a partner may provide a similar distraction.

Further, there were no differences in intention to exercise in the future or willingness to participate again in a similar study. Although SE was significantly correlated with persistence at the task, the motivation-enhancing effect of working with others was not mediated by changes in SE—there was no effect of experimental condition on SE ratings.

Our results did not show an additional boost in motivation above and beyond that produced by the social comparison due to the indispensability of our participants' efforts for the group score. That is, the conjunctive task demand condition did not produce significantly greater motivation than the additive or coaction conditions. Further, under the present conditions, working together in a team, under either additive or conjunctive task demands was not superior to simply exercising with a more capable coactive partner. Thus, the motivational gain observed here did not seem to depend on being interdependent, but simply required that participants were aware that they are being outperformed by a peer exercising on the same task at the same time.

At first blush, this is a bit puzzling because a number of previous studies have shown a significantly larger indispensability effect with conjunctive demands than with coactive ones (see Weber & Hertel, 2007 or Kerr & Hertel, 2011, for a review and meta-analysis). It is not plausible that our participants in the conjunctive condition simply misunderstood the interdependence imposed by their task demands; very similar instructions and procedures have been used in numerous previous studies that have produced robust indispensability effects, and in any case, large majorities of our participants correctly recalled the conjunctive (and additive) task demands in the postexperimental questionnaire. It is conceivable that the indispensability effect requires some minimal degree of group identification that was not achieved under the present experimental condition. Unfortunately, we did not assess whether participants in the conjunctive and additive conditions experienced a strong sense of group identity and membership, and whether they were committed to achieving a good group score. Still, it should be noted that several aspects of the present experimental procedures (e.g., seeing one's partner, a pretask verbal interaction, referring to the exercise group as a team) seem, if anything, more likely to foster group identification than the conditions typically used in many previous studies that have found such an indispensability effect (see Kerr & Hertel, 2011, for a review).

On the other hand, it has also been shown that when interpersonal competitiveness is high (e.g., via priming, Kerr et al., 2007, Exp. 3), even those who normally exhibit an indispensability effect (e.g., females) may not do so. It may be that exercising with a peer, particularly when there is no prior basis for social comparison, is an inherently competitive context. This would explain both the smaller indispensability and absence of gender effects in the current study (Kerr et al., 2007). In this vein, it is interesting to note that Lount, Kerr, Messé, Seok, and Park (2008) suggested and found evidence for the proposition that the social comparison mechanism attenuates with repeated group interaction—the goal of achieving a favorable social comparison might be abandoned as unachievable following repeated failures. This also suggests that we might find indispensability effects emerging if the exercise game was played repeatedly over time, creating more opportunity for in-group identification to develop. We are currently examining this possibility in a follow-up study using an aerobic exercise (viz., stationary bike riding).

In addition, we offered rewards as an incentive for good performance, which has been typical in previous empirical demonstrations of the Köhler effect (e.g., Kerr et al., 2007). Specifically, participants were told that the higher their scores, the better chance they had to win a semester-long membership at a campus fitness center. Our assumption was that such an extrinsic reward would enhance effort, particularly when one's teammate's chances of obtaining the reward closely depended upon that effort (i.e., in the conjunctive condition). However, there is considerable research demonstrating the demotivating properties of extrinsic motivations (e.g., Deci, Koestner, & Ryan, 1999). Moreover, if our participants placed little value on the prize (e.g., because they had little interest in using the weights and other heavy equipment available at the fitness centers), they may have felt little responsibility to help attain that prize.[5] One might examine this possibility by offering a prize likely to be highly valued by all (e.g., a large cash incentive). However, because such extrinsic rewards are rarely offered for pursuing health and fitness goals, another more promising approach would be to eliminate extrinsic rewards altogether. It is important to determine whether indispensability effects might emerge if one simply knew that the group's score and the quality of one's partner's workout largely depended upon one's level of effort, even if there was no additional extrinsic reward at stake. Again, we are in the process of examining the possible demotivating effects of extrinsic rewards in follow-up study (see Kerr, Feltz, & Irwin, 2010, for preliminary findings).

Our virtually presented partners were moderately more capable (i.e., persisted 1.4 times longer) than participants because a moderate discrepancy in ability between partners has been shown to be optimal for producing the motivation gain (Köhler, 1927; Messé et al., 2002). Messé et al. have reported that the Köhler motivation gain effect is smaller when one's more capable partner is either only slightly more capable or extremely more capable than oneself. Future research should examine the discrepancy in ability between partners in health video games where the trainers or models who demonstrate an exercise, set a pace, and/or encourage maximal effort are usually exceptionally fit in appearance and far superior from the target exerciser. Additional promising areas for investigation include the effects of encouragement from the partner, similarity in appearance and age of the partner, and the degree to which the partner is live versus completely virtual.

No study is without its limitations, and this study is no exception. We employed only a single type of exercise—an isometric plank exercise. We cannot say that our findings will generalize to other exergames that are more dynamic in nature. In addition, the game and interaction with a virtually presented partner was a relatively brief one-time experience (approximately 10 min). Future studies will examine more dynamic exergames that involve longer interactions and over multiple sessions. Moreover, it may well be that with stronger levels of group identification, a somewhat different pattern of results might emerge (e.g., a robust indispensability effect). Further, though not the major purpose of our study, SE was measured with a single item, rather than a series of hierarchical presented persistence times. However, the findings were still in line with previous studies that measured SE on persistence tasks.

Our study is the first to investigate the Köhler effect on motivation in health video games. None of the extant health video games have incorporated the critical design features based on the group dynamics of social comparison and the indispensability of effort for team success to motivate vigorous exercise. Our early results suggest that working out with virtually present, superior partners can improve persistence motivation on exercise game tasks. These findings provide a starting point to test additional features that have the potential to improve motivational gains in health video games.

Notes

1. Toward the end of this experiment, we included heart rate data during rest periods as well for quality control purposes (some participants were having difficulty remembering to start and stop the monitor for each exercise). This affected only 13 participants and there were no significant differences in the pattern of HR data between this subset of participants and those for whom HR was recorded only during exercise.

2. Participants following a revised procedure used at the end of the study (see Note 1) pressed the button once at the beginning of the trial and again at the end ($n = 13$).

3. It is plausible that participants may have counted the time between trials in their heads to give them an indication of their partner's actual performance. However, even an accurate estimation of the time between trials (i.e., 30 s) would not clearly indicate the actual performance of the partner, because the partner could have quit at any time within the 30-s period. Second, the threat that knowledge of one's partner's ability has to internal validity is that it may skew the perception of one's partner's relative ability, which we wanted to hold constant at "moderately more capable." Because there were no differences in perceived partner ability between experimental conditions, the manipulation was successful.

4. The equivalence of the coaction and conjunctive conditions was confirmed in a more detailed analysis examining the five exercises separately. A 5 (exercise) × 2 (work condition: coactive vs. conjunctive) × 2 (gender) ANCOVA with repeated measures on the first factor and Block 1 persistence scores as a covariate showed no work condition main effect, $F(1,96) = 1.95$, *ns,* or Exercise × Work Condition interaction effect, $F(4,384) = .84$, *ns.*

5. It is important to note that devaluing the prize need not undermine the goal of striving for a favorable social comparison with one's partner/teammate.

Acknowledgments

This research was supported by the Robert Wood Johnson Foundation. Special thanks to Joe Eisenmann for his assistance on design, and to Sara Schimpke, Matt Steck, Nate Keniston, Nik Skogsberg, Katherine Corker, Edward Witt, and Seunghyun Hwang for their assistance in conducting this study.

References

Bain, L.L., Wilson, T., & Chaikind, E. (1989). Participant perceptions of exercise programs for overweight women. *Research Quarterly for Exercise and Sport, 60,* 134–143.

Borg, G.A.V. (1998). *Borg's perceived exertion and pain scales*. Champaign, IL: Human Kinetics.

Carron, A.V., Hausenblas, H.A., & Mack, D. (1996). Social influence and exercise: A meta-analysis. *Journal of Sport & Exercise Psychology, 18,* 1–16.

Chatzisarantis, N.L.D., Hagger, M.S., Biddle, S.J.H., & Smith, B. (2005). The stability of the attitude-intention relationship in the context of physical activity. *Journal of Sports Sciences, 23,* 49–61.

Coleman, L., Cox, L., & Roker, D. (2008). Girls and young women's participation in physical activity: psychological and social influences. *Health Education Research, 23,* 633–647.

Deci, E.L., Koestner, R., & Ryan, R.M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125,* 627–668.

Dishman, R.K., & Buckworth, J. (1996). Increasing physical activity: A quantitative synthesis. *Medicine and Science in Sports and Exercise, 28,* 706–719.

Edwards, J.R., & Parry, M.E. (1993). On the use of polynomial regression equations as an alternative to difference scores in organizational research. *Academy of Management Journal, 36,* 1577–1613.

Feltz, D.L., & Riessinger, C.A. (1990). Effects on In vivo emotive imagery and performance feedback on self-efficacy and muscular endurance. *Journal of Sport & Exercise Psychology, 12,* 132–143.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7,* 117–140.

Fitzgerald, S.G., Cooper, R.A., Thorman, T., Cooper, R., Guo, S., & Boninger, M.L. (2004). The GAME(Cycle) exercise system: comparison with standard ergonometry. *The Journal of Spinal Cord Medicine, 27,* 453–459.

Fox, L.D., Rejeski, W.J., & Gauvin, L. (2000). Effect of leadership style and group dynamics on enjoyment of physical activity. *American Journal of Health Promotion, 14,* 277–283.

Franzini, L., Elliott, M.N., Cuccaro, P., Schuster, M., Gilliland, M.J., Grunbaum, J.A., et al. (2009). Influences of physical and social neighborhood environments on children's physical activity and obesity. *American Journal of Public Health, 99,* 271–278.

Haddock, B.L., Siegel, S.R., & Wikin, L.D. (2009). The addition of a video game to stationary cycling: The impact on energy expenditure in overweight children. *The Open Sports Sciences Journal, 2,* 42–46.

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.

Hertel, G., Kerr, N.L., Scheffler, M., Geister, S., & Messé, L.A. (2000). Instrumentality effects on motivation gains in groups: The role of impression management and spontaneous goal setting in promoting the Köhler effect. *Zeitschrift für Sozialpsychologie, 31,* 204–220.
Hertel, G., Niemeyer, G., & Clauss, A. (2008). Social indispensability or social comparison: The why and when of motivation gains of inferior group members. *Journal of Applied Social Psychology, 38,* 1329–1363.
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.
Kendzierski, D., & DeCarlo, K.J. (1991). Physical Activity Enjoyment Scale: Two validation studies. *Journal of Sport & Exercise Psychology, 13,* 50–64.
Kerr, N.L., Feltz, D.L., & Irwin, B.C. (2010, July). Explicit incentives and the Köhler motivation gain in exercise groups. Paper presented at the meeting of the Interdisciplinary Network for Group Research (INGRoup), Washington, D.C.
Kerr, N.L., & Hertel, G. (2011). The Köhler group motivation gain: How to motivate the "weak links" in groups. *Social and Personality Psychology Compass, 5,* 43–55.
Kerr, N.L., Messé, L.A., Seok, D.H., Sambolec, E.J., Lount, R.B., Jr., & Park, E.S. (2007). Psychological mechanisms underlying the Köhler motivation gain. *Personality and Social Psychology Bulletin, 33,* 828–841.
Köhler, O. (1927). Über den Gruppenwirkungsgrad der menschlichen Körperarbeit und die Bedingung optimaler Kollektivkraftreaktion [Human physical performance in groups and conditions for optimal collective performance]. *Industrielle Psychotechnik, 4,* 209–226.
Lount, R.B., Jr., Kerr, N.L., Messé, L.A., Seok, D.H., & Park, E.S. (2008). An examination of the stability and persistence of the Köhler motivation gain effect. *Group Dynamics: Theory, Research, and Practice, 12,* 279–289.
Lount, R.B., Jr., Park, E.S., Kerr, N.L., Messé, L.A., & Seok, D. (2008). Evaluation concerns and the Köhler Effect: The impact of physical presence on motivation gains. *Small Group Research, 39,* 795–812.
McArthur, L.H., & Raedeke, T.D. (2009). Race and sex differences in college student physical activity correlates. *American Journal of Health Behavior, 33,* 80–89.
McAuley, E. (1993). Self-efficacy and the maintenance of exercise participation in older adults. *Journal of Behavioral Medicine, 16,* 103–113.
Messé, L.A., Hertel, G., Kerr, N.L., Lount, R.B., Jr., & Park, E.S. (2002). Knowledge of partner's ability as a moderator of group motivation gains: An exploration of the Köhler discrepancy effect. *Journal of Personality and Social Psychology, 82,* 935–946.
Moritz, S.E., Feltz, D.L., Fahrbach, K., & Mack, D. (2000). The relation of self-efficacy measures to sport performance: A meta-analytic review. *Research Quarterly for Exercise and Sport, 71,* 280–294.
Porcari, J., Schmidt, K., & Foster, C. (2008). The Wii: As good as the real thing? *ACE FitnessMatters, July/August,* 7-9. Retrieved from http//www.acefitness.org/getfit/studies/wiistudy.pdf.
Raedeke, T.D. (2007). The relationship between enjoyment and affective responses to exercise. *Journal of Applied Sport Psychology, 19,* 105–115.
Rhodes, R.E., & Courneya, K.S. (2005). Threshold assessment of attitude, subjective norm, and perceived behavioral control for predicting exercise intention and behavior. *Psychology of Sport and Exercise, 6,* 349–361.
Sallis, J.F., Howell, M.F., Hofstetter, C.R., Elder, J.P., Hackley, M., Caspersen, C.J., et al. (1990). Distance between homes and exercise facilities related to frequency of exercise among San Diego residents. *Public Health Reports, 105,* 179–180.
Sallis, J.M., Prochaska, J.J., Taylor, W.C., Hill, J.O., & Geraci, J.C. (1999). Correlates of physical activity in a national sample of girls and boys in grades 4 through 12. *Health Psychology, 18,* 410–415.

Seta, J.J. (1982). The impact of comparison processes on coactors' task performance. *Journal of Personality and Social Psychology, 42,* 281–291.

Steiner, I.D. (1972). *Group process and productivity*. New York: Academic Press.

U.S. Department of Health and Human Services. (2008). *2008 Physical activity guidelines for Americans (ODPHP Publication No. U0036)*. Washington, DC: Physical Activity Guidelines Advisory Committee.

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: Lawrence Erlbaum.

Suls, J., & Wheeler, L. (2000). *Handbook of Social Comparison: Theory and Research*. New York: Kluwer Academic/Plenum Publishers.

Vroom, V.J. (1964). *Work and motivation*. New York: John Wiley.

Weber, B., & Hertel, G. (2007). Motivation gains of inferior group members: A meta-analytical review. *Journal of Personality and Social Psychology, 93,* 973–993.

Zakarian, J.M., Hovell, M.F., Hofstetter, C.R., Sallis, J.F., & Keating, K.J. (1994). Correlates of vigorous exercise in a predominantly low SES and minority high school population: A behavioral epidemiological study. *Preventive Medicine, 23,* 314–321.

Manuscript submitted: September 24, 2010

Revision accepted: March 24, 2011