"…been down so long…": perpetual vs. intermittent inferiority and the Köhler group motivation gain in exercise groups

Norbert L. Kerr, Samuel T. Forlenza, Brandon C. Irwin, Deborah L. Feltz

Published Version Information

Citation: Kerr, N. L., Forlenza, S. T., Irwin, B. C., & Feltz, D. L. (2013). "…been down so long…": Perpetual vs. intermittent inferiority and the Köhler group motivation gain in exercise groups. Group Dynamics: Theory, Research, and Practice, 17(2), 67-80.

Digital Object Identifier (DOI): doi:10.1037/a0031588

Publisher's Link: http://psycnet.apa.org/journals/gdn/17/2/67/

”…been down so long…”:

Perpetual vs. Intermittent Inferiority and the Köhler Group Motivation Gain in Exercise Groups

Norbert L. Kerr

Michigan State University and University of Kent

Samuel T. Forlenza

Michigan State University

Brandon C. Irwin

Kansas State University

Deborah L. Feltz

Michigan State University

Date: 11/15//2012 (PREPUBLICATION DRAFT. DO NOT CITE WITHOUT PERMISSION)

Word count: 8499

Author Note: Support for this research was provided by the Robert Wood Johnson Foundation’s Pioneer Portfolio through a grant from its national program, “Health Games Research: Advancing Effectiveness of Interactive Games for Health.” Portions of this paper were presented at the 2012 INGRoup Conference, Chicago, IL. Correspondence should be addressed to Norbert L. Kerr, Michigan State University, East Lansing, MI 48824. E-mail: kerr@msu.edu; Fax: (517) 353-1652.

ABSTRACT

Prior research has documented the Köhler motivation gain effect-- working with a more capable partner at a task that makes one's performance indispensible for the group can boost task motivation. Recent research has shown that the Köhler effect can boost one's persistence exercising in groups, but that always being the group's "weak link" can eventually undermine these motivation gains. An experiment is reported which contrasts having a partner that is more capable on all/both exercise tasks with one that is more capable on the focal task, but inferior on the second task. The Köhler effect on the focal task was replicated and unmoderated by the uniformity of the partner's exercise superiority. Implications for further research and application are discussed.

"…been down so long…": Perpetual vs. Intermittent Inferiority and the Köhler Group Motivation Gain in Exercise Groups

Most people recognize the value of exercise and many try to initiate an exercise program. However, it is notoriously hard to persist in keeping such good resolutions—e.g., research has typically reported about a 50% drop out rate within the first 6 months of initiating an exercise program (Dishman, 1994). Thus, a key question in any attempt to improve fitness is, "how can motivation to exercise be enhanced?" The present paper reports research that combines two promising approaches to answering this question—1) the use of exercise video games, or 'exergames' (e.g., Wii Fit, PlayStation 2 EyeToy: Kinetic) and 2) the application of social psychological principles that have been shown to enhance member motivation in work groups.

Clearly, video games have captured the attention of many people, particularly young people. For example, Gentile (2009) reports that 8-18 year olds in the U.S. spend about 12 - 14 hours per week, on average, playing video games. Such high levels of game play has raised a number of concerns (e.g., about the effects of graphic and violent video games, e.g., Anderson et al., 2010). One such concern is that video game play is typically very sedentary. But if engaging video games could be designed to require vigourous activity, such play could make a significant contribution to fitness. There are a number of obvious advantages to such games over alternative means of exercising—they can be played in the home at one's convenience, help avoid social physique anxiety, and provide timely performance feedback, etc. In the last decade, a number of such games have been designed (Staiano & Calvert, 2011). Although initial research efforts into the effectiveness of such games is encouraging (e.g., Maddison et al., 2007; Porcari, Schmidt, & Foster, 2008), as of yet there is rather little research exploring whether and when such games can succeed in boosting motivation to exercise. Moreover, practically none of the extant exergames incorporate social psychological principles which have been shown to boost motivation in group task settings.

Group Dynamics and Motivation to Exercise

Researchers have consistently found group exercise leads to higher exercise adherence than individual exercise programs (Burke, Carron, Eys, Ntoumanis, & Estabrooks, 2006; Dishman & Buckworth, 1996). Specifically, group exercise programs provide opportunities for comparison with others and are related to higher enjoyment and levels of social support, as well as increased intention to continue exercising. However, prior studies of group exercise rarely (if ever) introduce any real interdependence between exercisers (e.g., create teams whose progress and/or outcomes are mutually determined). A large and growing body of basic research on motivation in task groups (e.g., see Baron & Kerr, 2003; Karau & Williams, 1993) has demonstrated that certain patterns of interdependence can lead to group motivation gains (i.e., instances in which group members exert greater effort at a task than they would were they working individually). One particularly well-studied motivation gain phenomenon is the *Köhler effect.* The Köhler effect occurs when low ability group members increase their motivation 1) due to an invidious social comparison with their more capable group members and 2) when low ability group members see their contributions as particularly indispensible for group success and social evaluation (Kerr et al., 2007). Many factors that moderate this effect have also been identified (e.g., group sex composition, Lount et al., 2000; how long the group has worked together; Lount et al., 2008). Most importantly, the demands of the group task can be crucial; group tasks that make the least capable member's performance critical for group success (e.g., a mountain climbing team tethered together and able to climb at the speed of the slowest climber--what Steiner, 1972, called a *conjunctive group task*) generally results in the strongest Köhler effect (Hertel et al., 2008; Kerr et al., 2007) See Weber and Hertel (2007) for a meta-analytic or Kerr and Hertel (2011) for a narrative review of this literature.

The utility of the Köhler effect for enhancing motivation to exercise has been demonstrated in a series of recent studies from our lab. In the first (Feltz, Kerr, & Irwin, 2011) we found that exercising with a virtually-present, more capable partner led to a 24% improvement in persistence at a series of isometric plank exercises, relative to exercising alone. In the second (Irwin et al., in press), we found

that exercising with a virtually-present, more capable partner at a conjunctive team exercise task (i.e., when it was the first person to quit who defined the team's overall performance) led to a remarkable 125% improvement in persistence at an aerobic exercise task (viz., riding a stationary bike), again relative to exercising alone. In the third (Feltz, Irwin, & Kerr, 2012), we replicated the basic effect in a conjunctive version of the isometric exercise task (overall, an increase in persistence of 48%, relative to individual exercise). It seems clear that the Köhler effect can be effective in boosting motivation to exercise of the less fit member of an exercise group.

How "less fit"?

The question we consider in this paper is, "what's the optimal inferiority of the less fit member for producing the Köhler motivation gain?" At least one answer to this question has already been well documented—the Köhler motivation gain is maximal when the discrepancy in ability between the less and more capable group members is moderate in size (i.e., the ratio of the stronger's performance to the weaker's is ~1.4; cf. Messé et al., 2002; Köhler, 1926, 1927). In Köhler's original studies (Köhler, 1926, 1927), more recent lab research (e.g., Messé et al., 2002), and in our own recent exercise research (Feltz et al., 2012), results have shown that the relationship between discrepancy in teammates' abilities and the gain in motivation showed by the weaker member is a curvilnear one. For purposes of illustration, Messé et al.'s results are plotted in Figure 1, in both an experiment in which dyads were randomly composed and the discrepancy in ability could take on any value (Figure 1a) and another in which this discrepancy was experimentally manipulated (Figure 1b). An explanation for this inverted-U function has been provided by Kerr, Seok, & Messé (2007). They manipulated both the discrepancy in ability (as in Messé et al.'s, 2002, Exp. 2) and the interdependence between the co-workers. In one condition (Additive), the two workers could observe one another's performance (at a taxing arm-lifting task), and the group's score was the simple sum of member scores. This condition permits social comparison, but neither member's contributions were particularly indispensable for the other's outcomes. In the other condition (Conjunctive), not only could the two workers observe one another, but they worked as a team

under conjunctive task demands (i.e., the group's score would be defined by the member who quit the arm-lifting task first). In this condition, both of the processes that can contribute to the Köhler effect (i.e., invidious social comparison; feeling indispensible to the group) could operate fully. Kerr et al.'s (2007) results are presented in Figure 2. The inverted-U function is evident in both conditions, but motivation (here, Block 2 - Block 1 task persistence, where Block 1 is a performance baseline collected before experimental manipulations) is also significantly higher in the conjunctive than the additive condition. From these results, Kerr et al. (2007) concluded that partner discrepancy did not moderate the indispensibility mechanism (i.e., regardless of how much better my partner is than me, if I'm the group's "weak link," I realize that my performance is crucial for the group's outcome under conjunctive task demands), but that it did moderate the social comparison mechanism. As Festinger (1954) suggested in the original statement of social comparison theory, in the realm of abilities, one is less likely to compare with someone of about the same ability (low discrepancy) than with someone of somewhat higher ability (moderate discrepancy); the former is not particularly informative, but the latter can present a real challenge. But in addition, as has been noted in much goal-setting research (e.g., Locke, 2001), a goal (e.g., matching one's partner's performance) has to be seen as achievable to be motivating, and if one's partner is too much more capable (high discrepancy), one may simply stop comparing oneself to that partner (i.e., give up).

Another route to the belief that it is impossible to compare favorably with one's partner is consistent failure to match or exceed a partner's performance. Lount et al. (2008) found that working with a single partner who was repeatedly more capable at a conjunctive group task led to smaller motivation gains than working with a series of different partners who all were, likewise, repeatedly more capable. They suggested that people are less likely to become resigned to the role of the perpetually inferior group member in the latter case—with each new partner comes the possibility that one could compare more favorably. So, one way in which to maintain a robust Köhler motivation gain across repeated exercise sessions might be to introduce new partners across sessions. But this may be less

practical in a team exercise setting than having a partner who is more capable than the player on some exercises, but less capable on others. As in the new-partner condition of the Lount et al. (2008) study, this avoids putting a player in the discouraging role of the perpetually inferior group member.

In the present experiment, we address the question, "will the duration of exercise be greater when one is for at least some exercises not his/her exercise team's 'weak link,' compared to when s/he is the 'weak link' for all exercises?" If the answer is yes, this would prescribe that exergames use a variety of exercises (already a good idea for approximating a full-body workout), but also that for some of these exercises, the game player is more capable than his/her teammate. This could be achieved by careful matching of teammates and exercise tasks, or more easily in games with genuinely virtual teammates, by simply having that virtual teammate be less capable than the player for some of the game's exercises.

Overview of the Experiment

Participants performed four blocks of exercise trials, with two exercises per block. The two exercises were 1) an isometric plank exercise (the participant held a position much like a push-up, face down on a cushioned mat, with legs extended straight, and the body lifted upward by resting the forearms and toes on the mat; the body, spine, and legs were in a straight line and nothing was touching the ground except for the forearms and toes) and 2) a wall-sit exercise (squatting into a sitting position with one's back against a wall). The participants' goal was to hold each exercise for as long as possible. The first block provided a baseline measure which was used to control for individual differences in task ability and task motivation. In individual-control conditions, participants worked individually at each of the four trials. This condition provided an estimate of fatigue and boredom effects across Blocks 2-4. In the remaining experimental conditions, participants were introduced to a same-sex teammate after the first block and told that they would be working as a two-person team on all subsequent blocks. Further, the dyad's task demands were conjunctive—the trial ended as soon as one dyad member quit—and the time that member persisted served as the group's score on that trial.

The primary experimental manipulation consisted of feedback indicating either that the participant's teammate was moderately more capable at both of the exercises (S-always-inferior conditions), or that the teammate was moderately more capable at only one of the two exercises (S-sometimes-superior conditions). Half of the latter conditions had the teammate inferior on the plank exercise; for the other half, the teammate was inferior on the wall-sit exercise. Much prior research suggests that participants will persist longer in the S-always-inferior condition than in the individual-control condition—that is, that there will be a significant Köhler motivation gain. The results of Lount et al.'s (2008) study, with a roughly comparable design, suggest that the magnitude of this effect will attenuate across trials (one tends to become discouraged and give up on keeping up with a consistently superior partner). The interesting and open research question is whether this effect will be stronger and persist longer when the participant is not always inferior to his/her partner.

Method

Design and Participants

Participants were 114 undergraduate students who completed the experiment in return for course credit. Seven participants who expressed suspicion about whether their partner was genuine were dropped from the analyses, so the final sample consisted of 107 participants (52 males, 55 females; *M Age* = 19.2 years, *SD* = 1.18). Within each gender, participants were randomly assigned to one of eight conditions (1: individual-control, plank exercise first; 2: individual-control, wall-sit exercise first; 3: partner superior on both exercises, plank first; 4: partner superior on both exercises, wall-sit first; 5: partner superior only on plank exercise, plank first; 6: partner superior only on plank exercise, wall-sit first; 7: partner superior only on wall-sit exercise, plank first; and 8: partner superior only on wall-sit exercise, wall-sit first). For purposes of description, one could frame the design as a 2 (Gender) x 4 (Condition: Individual-control, S-always-inferior, S-superior on plank, S-superior on wall-sit) x 2 (Exercise order: Plank first, Wall-sit first) x 4 (Block) x 2 (Exercise: Plank, Wall-sit) factorial with

repeated measures on the last two factors. However, the comparisons of interest will focus on subsets of these conditions.

Procedure

Participants arrived individually. After signing an informed consent, both exercises were demonstrated. The plank exercise was demonstrated first with the use of the EyeToy: Kinetic game on the PlayStation 2. EyeToy: Kinetic is an exercise video game that allows the player to perform various types of physical activity (e.g., yoga, cardio exercises, combative exercises) with the assistance of a virtual trainer. Included with the EyeToy: Kinetic game is a video camera that projects the player's image onto the television screen, which allows the player to interact with the virtual environments the game creates. The wall-sit exercise was demonstrated second by one of the experimenters on a pre-recorded video because the EyeToy: Kinetic game did not have the wall-sit exercise programmed into it.

All participants then individually completed the two exercises with a short (40s) rest between exercises; this constituted the first block of exercises (Block 1). Roughly half of all participants did the plank exercise first (on this and all succeeding blocks); the remainder did the wall-sit first. After a rest period, participants in the individual-control conditions were told how long they held each exercise and that they would be repeating the same block of exercises three more times individually.

After Block 1, participants in the remaining experimental conditions were told they would be repeating the exercises with a same-sex partner who was connected to the lab through the internet. As in the Feltz et al. (2011) study, the participants were introduced to their virtual partner over a Skype-like connection. The partner was presented as a similar-aged college student dressed in workout clothes and with average height, weight, and build. Although participants thought they were interacting with a live person, in reality, their partner was a confederate whose Skype and exercise performance videos were both pre-recorded. While exercising during the remaining trials, participants could see their teammate's image on a screen alongside their own image and were told their partner could likewise see their image on a screen.

After the Skype introduction, participants were told accurately how long they had actually held each exercise and given false feedback about how long their partner had held each exercise during the first block. The purported length of time their partner held the exercises depended on which condition the participant was in. As noted above, previous research has established that the Köhler motivation gain is maximal when one's partner is moderately more capable (e.g., Feltz et al., 2012). Therefore, participants who had a partner that was superior on both exercises were told their partner held both the plank and wall-sit for a moderately longer length of time (i.e., 40% longer than the participants had done, a discrepancy that prior research suggests produces a robust Köhler effect; Messé et al., 2002; Köhler, 1926, 1927). Participants who had a partner that was superior on only one exercise were told their partner held either the plank or the wall-sit for moderately (i.e., 40%) longer than s/he had done, while the participant allegedly held the other exercise for moderately longer time (viz., the partner persistence was 40% less than the participant's. In this way, participants were led to believe they were better than their partner on either one exercise or neither exercise.

Experimental participants were also told that performance on the coming trials would be measured using a team score. The team score was defined by the persistence time of the person who quit first. Therefore, when one person stopped an exercise, the other person had to stop exercising, and the team's score was however long the first person lasted.

During exercises where the teammate was allegedly better than the participant, the teammate's performance always lasted longer than the participant's. This was achieved with a looping video showing the partner holding his/her exercise and not quitting until after the participant did. In these cases, the participant was not told how much longer the allegedly-more-capable partner did or could have persisted; rather, consistent with the conjunctive task demands, as soon as the participant quit, the experimenter communicated this to the other lab, instructing the (fictional) experimenter there to tell the partner to quit. On exercises where the participant was allegedly better than his/her teammate, the teammate's performance was manipulated to be inferior to the participant's. This was accomplished by

the experimenter visually monitoring the participant and when it appeared that the participant was close to quitting, the experimenter would tell participants that their partner stopped (the feed of the partner's image was paused when they quit) and therefore they could stop. Thus, participants were either continually outperformed by their partner on both exercises or were able to outperform their partner on one of the exercises.

Measures

Persistence. Level of task motivation was assessed from persistence at the task. Persistence was the total number of seconds an exercise was held by participants from the moment they were in position to the moment they quit holding the exercise.

Ratings of perceived exertion (RPE). Perceived exertion was measured using the Borg RPE scale (Borg, 1998). The scale ranges from 6 - 20 where 6 means "no exertion at all" and 20 means "maximal exertion." Participants were asked to rate their total feeling of exertion at the end of each exercise, with particular reference to their perceived exertion immediately before the end of the exercise.

Self-efficacy beliefs (SE). SE was measured at three points during the experiment: before the first block, before the second block, and after all exercises were completed. Participants were asked how many seconds they believed they could hold each exercise. The second measurement (before the second block) occurred directly after participants were told how long they held each exercise during the first block. Additionally, participants in the team conditions were already introduced to their partners and knew how long their partner held each exercise.

Enjoyment and intention to exercise in the future. Enjoyment of the activity and intention to exercise were measured at the end of the fourth trial block. Enjoyment was measured using a short version (8-items) of the Physical Activity Enjoyment Scale (PAES; Kendzierski & DeCarlo, 1991). Each item was rated on a 7-point bipolar scale regarding the enjoyment, interest, pleasantness, etc. of the activity. Intention to exercise was measured with one-item that asked 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"). Validity of a single item in measuring this construct has been demonstrated within the exercise domain (Chatzisarantis, Hagger, Biddle & Smith, 2005; Rhodes & Courneya, 2005).

Results

Exercise Persistence

Following the analyses conducted in earlier studies, exercise persistence was taken as the time participants held each exercise within each block. There are, of course, individual differences in fitness and strength that we wished to control for. This could be done in different ways. In many prior studies (e.g., Hertel et al., 2000; Kerr et al., 2007), participants' Block 1 performance was used as a baseline and the difference score (e.g., Block 2 – Block 1) was the primary dependent variable, one that expressed each participant's Block 2 score relative to his/her Block 1 score (which, of course, reflected individual differences in fitness and strength). An alternative, less vulnerable to certain problems that can 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 later blocks' scores. Here, we present the results using the former, difference-score method because the mean values presented for such an analysis are more easily interpreted, and because they simplify the initial phase of the analysis (see the next paragraph). However, the two methods produced the same pattern of results for all contrasts of interest (viz., those involving the Condition factor). The actual persistence means which were the basis for the difference scores can be obtained (across blocks and conditions) at www.msu.edu/user/kerr/been-down-so-long/Means.pdf .

The initial analysis sought to replicate prior research (Feltz et al., 2011; Irwin et al., 2012) by contrasting individual exercisers (individual-controls) with those who believed they were always the inferior member of an exercise dyad working under conjunctive task demands (i.e., the dyad's score was determined by the poorer performance in the dyad; these were our S-always-inferior participants). For each exercise, a difference score between later and first block performance was computed (i.e., Plank.Block 2Δ = Block 2 Plank persistence – Block 1 Plank persistence, Plank.Block 3Δ = Block 3

Plank persistence – Block 1 Plank persistence, … , Wall-sit.Block 4Δ = Block 4 Wall-sit persistence – Block 1 Wall-sit persistence). These scores were then analyzed in a 2 (Gender) x 2 (Order: Plank first vs. Wall-sit first) x 2 (Condition: Individuals vs. S-always-inferior) x 3 (Block) x 2 (Exercise: Plank and Wall-sit) ANOVA with repeated measures on the last two factors. Because the Mauchly's test of sphericity indicated violation of the equality of variance of difference scores, the most conservative, Lower-bound correction was applied to all tests involving repeated measures factors (Girden, 1992).

Two of the significant effects were not of particular interest: 1) performance dropped across blocks, due to fatigue (Block 2 = 9.54s, Block 3 = -20.19s, Block 4 = -21.06s; $F(1, 50) = 49.94$, $p < .001$, $\eta_p^2 = .50$), and 2) relative to baseline, overall performance was better for females than for males (Males = -18.51s, Females = -2.63s; $F(1,50) = 7.51$, $p < .01$, $\eta_p^2 = .13$).[1] Of particular interest were the effects involving the individual vs. dyad contrast. First, there was a significant Köhler motivation gain: Individuals' M = -23.87s, S-always-inferior = 2.73s; $F(1,50) = 21.05$, $p < .001$, $\eta_p^2 = .296$. Second, as suggested by Lount et al.'s (2008) findings, this group motivation gain tended to attenuate across trials, Block x Condition, $F(1,50) = 3.99$, $p = .051$[2], $\eta_p^2 = .074$. The latter interaction is illustrated in Figure 3. No other effects were significant. Thus, these patterns did not depend upon which exercise was done, which was done first, or the sex of the participant.

Our remaining research question requires the comparison of the S-always-inferior condition to the S-sometimes-superior conditions. However, unlike the previous, omnibus analysis that included both exercise tasks, the present analysis requires that we look at each task separately. This is because we are not interested in performance at the task where one's partner appears to be inferior; by design, for these trials on Blocks 2 - 4 the partner quit as soon as the actual participant showed any signs of fatigue, making these untaxing trials uninformative about the participants' willingness to exert maximal effort. Rather, we want to focus on comparable conditions—where one's partner was alleged to be superior at the task and always persisted longer than the actual participant. So, for the Plank exercise, this would mean comparing plank exercise performance between the "S-always-inferior" and the "S-superior on

wall-sit" conditions. For the Wall-sit exercise, it would mean comparing wall-sit performance between the "S-always-inferior" and the "S-superior on plank" conditions.

The first of these analyses was a 2 (Gender) x 2 (Order) x 2 (Condition: S-always-inferior vs. S-superior on wall-sit) x 3 (Block) ANOVA on plank exercise trials only, with repeated measures on the final factor. Because the Mauchly's test of sphericity indicated violation of the equality of variance of difference scores, the most conservative, Lower-bound correction was again applied to all tests involving repeated measures factors (Girden, 1992). Two effects emerged: 1) the performance drop on later trials due to fatigue, Block $F(1,42) = 93.60$, $p < .001$, $\eta_p^2 = .69$ (see Figure 4a), and 2) a Block x Gender interaction, indicating that the relative advantage of females-to-males increased with blocks, $F(1,42) = 8.67$, $p < .01$, $\eta_p^2 = .171$ (see footnote 1, *Supra*). However, there were no significant effects involving the Condition factor— when one was inferior to one's partner on the plank exercise, it did not matter whether one was also inferior vs. superior to that partner on a second (wall-sit) exercise (e.g., see Figure 4a).

A parallel analysis was run on wall-sit trials only, using a 2 (Gender) x 2 (Order) x 2 (Condition: S-always-inferior vs. S-superior on plank) x 3 (Block) ANOVA with repeated measures on the final factor, and the conservative Lower-bound correction applied. The only significant effect was the fatigue-driven Block factor, $F(1,44) = 9.60$, $p < .001$, $\eta_p^2 = .179$.[3] As for the plank exercise, it made no difference for wall-sit performance whether one's partner was or was not superior on the other exercise. As Figure 4 shows, there are hints in late blocks of superior performance in the mixed-ability conditions, but these are not statistically significant, nor are any other effects.

Ancillary Analyses

The contrast of individuals with dyads in which the partner was always superior in the present experiment parallels in most respects the design used in Feltz et al. (2011). As noted above, we replicated the robust Köhler effect reported in that and several other recent studies. Feltz et al. summarized their remaining findings as follows: "The motivation gains achieved with a more capable

partner did not come at the expense of aversion to the task. … there was no evidence that they perceived they were working any harder, enjoyed the exercise less, or had lower [self-efficacy] SE at the task than controls. Further, there were no differences in intention to exercise in the future or willingness to participate again in a similar study" (p. 521). Our remaining analyses for the present study 1) sought to see whether this pattern of null effects were replicated here, and 2) whether sometimes being superior to one's partner had any different effect on these variables.

Report of physical exertion (RPE). First, in the initial analysis of RPE and in subsequent ancillary dependent variables—those that replicate the conditions of prior research most closely--we contrasted individual exercisers (our individual-controls) with those who believed they were always the inferior member of an exercise dyad working under conjunctive task demands (i.e., the dyad's score was determined by the poorer performance in the dyad). RPE measures were collected immediately after each exercise was completed. These were analyzed in a 2 (Condition: Individuals vs. Dyad/S-always-inferior) x 2 (Gender) x 2 (Order: Plank first vs. Wall-sit first) x 2 (Exercise) x 4 (Block) ANOVA with repeated measures on the last two factors. In light of a significant Mauchly test, the conservative Lower-bound correction was applied. The only significant effect to emerge was for Block, $F(1,150) = 30.46$, $p < .001$, $\eta_p^2 = .38$; unsurprisingly, fatigue led participants to report greater subjective effort on later trial blocks (Block 1 = 12.86, Block 2 = 13.93, Block 3 = 13.74, Block 4 = 13.61). But even though participants persisted ~27s longer in the dyad conditions, there was no significant effect of condition on subjective effort, either overall or moderated by block.

Second, as with the performance analyses, when analyzing RPE scores for the S-sometimes-superior conditions, we needed to do separate analyses on the two exercises. The appropriate analysis on the RPE scores for the plank exercise [viz., a 3 (Condition: Individual controls vs. S-always-inferior vs. S-superior on wall-sit) x 2 (Gender) x 2 (Order) x 4 (Block) repeated-measures ANOVA] again only resulted in one significant effect—the fatigue-driven Block main effect, $F(1,207) = 25.73$, $p < .001$, $\eta_p^2 = .27$. There was no effect of condition on subjective effort, either overall or moderated by block.

Planned contrasts showed that there was no difference between individual-controls and members of dyads where the participant was always inferior (p = .71), and further that there was no difference between the latter condition and dyads where participant was superior on the wall-sit (p = .13). The corresponding analysis for the wall-sit exercise [now with a 3 (Condition: Individual-controls vs. S-always-inferior vs. S-superior on plank) factor] showed the same pattern: a clear Block effect (F(1,213) = 39.06, p < .001, η_p^2 = .36), but no hint of a condition effect (all p's > .15 for the condition main effect, condition x block interaction, and planned contrasts between successive condition means).

After all trials were over, participants were also asked to rate how much effort they expended on all the prior exercises. An omnibus 4 (Condition) x 2 (Gender) x 2 (Order) between-Ss ANOVA confirmed the null pattern of the prior analyses. There was no significant effect of condition (p = .57), nor were planned contrasts significant [i.e., between the individuals and partner-always superior condition, or between the latter and the two participant-sometimes-inferior conditions, or between the latter two conditions (all p's > .39)].

Self-efficacy (SE). Because the initial SE judgment was collected prior to any experience with the task, it is best viewed as an index of chronic exercise SE. This recommended doing separate analyses on each exercise, so that the initial rating could be used as a covariate to control for chronic individual differences in SE. As in the preceding paragraph, one of the partner-sometimes-inferior conditions was dropped for each of these analyses. The first analysis, on the plank self-efficacy data, was a 3 (Condition: Individual controls vs. S-always-inferior vs. S-superior on wall-sit) x 2 (Gender) x 2 (Order) x 2 (Block: before Block 2 & after Block 4) with repeated measures on the last factor and the initial/before Block 1 plank self-efficacy score used as a covariate. Our primary interest was in determining whether effects involving the condition factor were significant. The condition main effect was not significant (p = .35), but the Condition x Block interaction was significant, F(2,68) = 3.14, p = .05, η_p^2 = .084. This was due to a condition main effect (p = .036) prior to Block 2; when participants had just been told that they had been outperformed by their partner on both exercises, they optimistically

predicted that they could persist significantly longer at the plank exercise (adjusted M = 80.21s) than when either they had received no partner feedback (individual-controls adjusted M = 61.59s) or had been told that they had only been outperformed by their partner on the wall-sit exercise (S-superior on wall-sit = 64.08s). However, by the end of the experiment, there was no hint of any such condition effect (p = .94).[4]

The second analysis, on the wall-sit SE data, was parallel to the first: a 3 (Condition: Individual controls vs. S-always-inferior vs. S-superior on plank) x 2 (Gender) x 2 (Order) x 2 (Block: before Block 2 & after Block 4) with repeated measures on the last factor and the initial/before Block 1 wall-sit self-efficacy score used as a covariate. Here there were no significant effects, including on the key condition factor (p = .35), for which none of the planned contrasts were significant (e.g., Individuals vs. S-always inferior, the latter vs. S-superior on plank; all p's > .10).

Intention to exercise. An omnibus 4 (Condition) x 2 (Gender) x 2 (Order) between-Ss ANOVA on participant's ratings of their "intention to exercise for at least 30 minutes the following day" showed no significant effects, and in particular, there was clearly no effect for experimental condition (p > .61). The grand mean was 1.58 (SD = 1.64) 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.

Enjoyment of the task. An overall enjoyment index was computed based on the 8 of the PAES scale. An omnibus 4 (Condition) x 2 (Gender) x 2 (Order) between-Ss ANOVA on this index yielded only a marginally significant condition effect, $F(3,90) = 2.61$, p = .056, η_p^2 = .08. Post-hoc Dunnette tests showed that participants enjoyed the task less when they had a partner who consistently outperformed them on both exercises (S-always-inferior M = 3.86) than when they had no partner (individual-controls M = 4.39), but if they had managed to outperform their partner on either the plank exercise (M = 4.14) or on the wall-sit exercise (M = 4.24), they enjoyed the task as much as individual-controls (and, as shown by a planned contrast, marginally (p = .061) more than those in the S-always-

inferior condition). The same general pattern obtained for a similar measure ("how difficult did you find the exercises?"), but the differences were far smaller and not significant (condition main effect $p > .25$).

Discussion

As has been shown in several prior Köhler effect studies, we observed that working out with a moderately more capable exercise team partner under conjunctive task demands boosted one's effort relative to working out individually. Under the current experimental conditions, such dyadic participants persisted an average of 26.1s longer at the plank exercise, a 35.6% improvement over individual controls, and 29.6s longer at the wall-sit exercise, a 48.7% improvement over individual controls. Also, as observed in at least one prior study (Lount et al., 2008), the longer one worked with that more capable partner, the smaller the motivation gain (see Figure 3), an effect that can reasonably be attributed to mounting discouragement about keeping up with the superior partner. The novel question addressed in this experiment was whether the Köhler motivation gain would be altered if one had less to be discouraged about—that is, if one were actually superior to one's partner on a second task. Although participants did tend to enjoy working at the exercise tasks more when they consistently outperformed their partner on one of the tasks (vs. being outperformed by their partner on both tasks), this enjoyment was not accompanied by any stronger (or weaker) Köhler effect for that task at which they were inferior. Our participants seemed to approach each task independently of their relative ability on the other task. Also, as found in comparable previous studies, the greater motivation observed in the dyads was not accompanied by greater perceived effort, altered task self efficacy, or reduced intention to exercise in the future. This bodes well for translating such short term motivation gains into more enduring improvements in training and fitness.

The clearest implication of our findings is that the Köhler motivation gain is quite robust. To our knowledge, the basic effect has failed to emerge only under fairly unusual conditions—specifically, when no one, including the experimenter/supervisor, can detect which group member is inferior (Kerr et al., 2005); for females only, when their superior partner had previously ostracised them (Kerr et al.,

2008); or when there are large partner age differences for a task where such age differences might imply large ability differences as well (Seok, 2004). Meta-analyses of this literature (Weber & Hertel, 2007) further document the robustness of the effect. In the realm of exercise groups or exergames, this suggests that the principles underlying the Köhler effect may be applied with success across a reasonably wide range of conditions. The present experiment, in particular, suggests that the effect will occur as long as the conditions are appropriate for a target task (namely., moderate inferiority to one's teammate at a conjunctive group task), regardless of whether these conditions hold for other tasks (e.g., one is otherwise superior to one's partner).

This robustness to partner characteristics is complemented by another advantage of using the Köhler effect to boost motivation to exercise—it is not accompanied by aversion to the exercise task—people work harder but they do not report less enjoyment of the immediate exercise task, nor any drop in intention to exercise in the future. The conditions of the Köhler effect, although frustrating in certain regards, appear to create a challenge which people strive to meet and take some satisfaction in meeting (Hertel et al., 2000b). This contrasts with the other best documented group motivation gain, the *social compensation* effect (e.g., Karau et al., 2000; Williams & Karau, 1991). In the latter, one's partner is relatively incapable or unmotivated at the group's additive task, and the only way to insure group success is to work extra hard to compensate for that underperforming partner. Doing more than one's own share of the group work is normally seen as inequitable and aversive, and can actually result in motivation losses (Kerr, 1983). It is only when success at the task is seen as extremely important that such an inequitable arrangement can prompt social compensation rather than social loafing (Williams & Karau, 1991). Nevertheless, successful social compensation seems much more likely to produce task or group aversion than success via the Köhler effect. A final advantage of the Köhler effect is its demonstrated magnitude in exercise groups. Here, we observed improvements over individual exercise of between ~35% - 45%. In our previous research, the improvements have ranged between ~25% and

~125%. These gains compare quite favorably to the effects typically reported in physical activity intervention studies (Williams & French, 2011).

Like any study, the present one has limitations. The participants were not people in special need of improved fitness (e.g., morbidly obese or rehabilitation patients), but relatively healthy and active students. False feedback of partner ability was used and this may be impractical or unethical in some exercise settings. But some of the potential limitations have been addressed in other studies. For example, here the exergame and interaction with a virtually-presented partner was a relatively brief, one-time experience with a pair of persistence exercises. However, in Irwin et al. (2012), even stronger Köhler motivation gains were observed in a series of workouts over several days using a challenging aerobic exercise task. Still, it is quite possible that it takes more time for the effects of mixed- vs. uniform-inferiority to manifest themselves than was available in the present study (a one-hour session). It is interesting to note that the non-significant trends in the data were for higher motivation gains in the mixed-inferiority conditions across trial blocks (see Figures 4a & 4b). Given sufficient time, it is possible that these weak trends would become detectable, an interesting question for future research.

We noted earlier that one potential limitation of the Köhler effect for increasing task motivation in groups in general, and in exercise groups in particular, is that it can become discouraging to perpetually be the group's least capable member. The present study's results found only suggestive evidence to support one means of avoiding such discouragement—viz., to "mix-up" the capabilities of group members, so that one is the least capable member for some, but not all the group's activities. Future research should not only further explore this remedy, but alternative remedies. One such promising remedy is varying performance feedback so that increases in one's effort result in reducing (although not fully eliminating) the discrepancy between oneself and one's superior partner. Perhaps one need not regularly or even ever surpass one's partner to avoid discouragement so long as one perceives that there is some, perhaps gradual, improvement in one's relative standing in the group (cf. Jones et al.,

1968). Prior work with the Köhler effect has not systematically varied the discrepancy between the inferior and superior partner across time; future research should do so.

References

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., . . . Saleem, M. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: A meta-analytic review. *Psychological Bulletin, 136*(2), 151-173.

Baron, R. S., & Kerr, N. L. (2003). *Group process, group decision, group action* (2nd ed.). Buckingham, UK: Open University Press.

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

Burke, S.M., Carron, A.V., Eys, M.A., Ntoumanis, N., & Estabrooks, P.A. (2006). Group versus individual approach? A meta-analysis of the effectiveness of interventions to promote physical activity. *Sport & Exercise Psychology Review*, *2*(1), 13-29.

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.

Dishman, R. K. (1994). *Advances in exercise adherence*. Champaign, IL: Human Kinetics.

Dishman, R. K., & Buckworth, J. (1996). Increasing physical activity: A quantitative synthesis. *Medicine and Science in Sport 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., Irwin, B., & Kerr, N. L. (2012). Two-player partnered exergame for obesity prevention: Using discrepancy in players' ability as a strategy to motivate physical activity.. *Journal of Diabetes Science and Technology. 6*, 1-8.

Feltz, D. L., Kerr, N. L., & Irwin, B. (2011). Buddy up: The Köhler Effect applied to health games. *Journal of Sport and Exercise Psychology, 33(4),* 506-526.

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

Gentile, D. (2009). Pathological video-game use among youth ages 8 to 18: A national study. *Psychological Science, 20*, 594-602.

Girden, E.R. (1992). *ANOVA: Repeated measures.* Newbury Park, CA: Sage.

Hertel, G., Kerr, N. L., Messé, L. A. (2000). Motivation gains in groups: Paradigmatic and theoretical advances 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. (2000b). 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.

Irwin, B. C., Scorniaenchi, J., Kerr, N. L., Eisenmann, J. C., & Feltz, D. L. (2012). Examining the Köhler motivation effect in aerobic exercise. *Annals of Behavioral Medicine*. (DOI) 10.1007/s12160-012-9367-4

Jones, E. E., Rock, L., Shaver, K. G., Goethals, G. R., & Ward, L. M. (1968). Pattern of performance and ability attribution: An unexpected primacy effect. *Journal of Personality and Social Psychology, 10*(4), 317-340. doi:10.1037/h0026818

Karau, S. J., Markus, M. J., & Williams, K. D. (2000). On the elusive search for motivation gains in groups: Insights from the collective effort model. *Zeitschrift Für Sozialpsychologie, 31*(4), 179-190.

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 and Exercise Psychology*, *13*, 50-64.

Kerr, N. L. (1983). Motivation losses in task-performing groups: A social dilemma analysis. *Journal of Personality and Social Psychology, 45,* 819-828.

Kerr, N.L. (2006). *The experimental study of group motivation gains: Overview and critique of the study of the Köhler effect*. Paper presented at the 1st Annual INGRoup Conference, Pittsburgh, PA, July 29.

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. M., Park, E. S., & Sambolec, E. (2005). Identifiably, performance feedback and the Köhler effect. *Group Processes and Intergroup Relations, 8*, 375-390.

Kerr, N. L., Messé, L. M., Seok, D., Sambolec, E., Lount, R. B., & Park, E. S. (2007). Psychological mechanisms underlying the Köhler motivation gain. *Personality and Social Psychology Bulletin, 33*, 828-841.

Kerr, N. L., Seok, D., & Messé, L. M. (2007). *Moderators of the Köhler discrepancy effect*. Paper presented at the 2nd Annual INGRoup Conference, July 14, 2007, East Lansing, MI.

Kerr, N. L., Seok, D., Poulsen, J., Harris, D., & Messé, L. M. (2008). Social ostracism and group motivation gain. *European Journal of Social Psychology, 38*(4), 736-746.

Köhler, O. (1926). Kraftleistungen bei Einzel- und Gruppenarbeit [Physical performance in individual and group work]. *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.

Locke, E. A. (2001) "Motivation by Goal Setting," *Handbook of Organizational Behavior*, 2: 43-54.

Lount, R. B., Kerr, N. L., Messé, L. M., Seok, D., & 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., Messé, L. A., & Kerr, N. L. (2000). Trying harder for different reasons: Conjunctivity and sex composition as bases for motivation gains in performing groups. *Zeitschrift für Socialpsychologie, 31*, 221-230.

Maddison, R., Mhurchu, C. N., Jull, A., Jiang, Y., Prapavessis, H., & Rodgers, A. (2007). Energy expended playing video console games: An opportunity to increase children's physical activity? *Pediatric Exercise Science, 19*, 334-343.

Messé, L. A., Hertel, G., Kerr, N. L., Lount, R. B., & 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.

Porcari, J., Schmidt, K., & Foster, C. (2008). ACE-Sponsored research as reported in Anders, M., *ACE Fitness Matters*.

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.

Seok, D-H (2004). *Exploring self-efficacy as a possible moderator of the Köhler discrepancy effect.* (Master's thesis). Michigan State University.

Staiano, A. E. & Calvert, S. L. (2011). Exergames for physical education courses: Physical, social, and cognitive benefits. *Child Development Perspectives, 5*, 93-98.

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

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

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.

Williams S. L., & French, D. P. (2011). What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour- and are they the same? *Health Education Research, 26*(2), 308-22.

Footnotes

1. The latter gender difference tended to be larger in later trials, $F(1,50) = 3.67$, $p = .061$; the female-to-male superiority was 5.1s at Block 2, 20.0s at Block 3, and 22.5s at Block 4.
2. Using less conservative corrections, such as the Greenhouse-Geisser correction, this interaction was significant ($p = .038$) at the conventional alpha level.
3. The Block x Gender interaction observed for the plank data was not significant for the wall-sit data ($p = .101$).
4. There was one other significant effect--a Block x Order interaction, $F(1,68) = 5.30$, $p = .024$, $\eta_p^2 = .072$; the order of working on the tasks had no effect when participants were about to begin the key trials, but by the end of the study, those who always began each trial with the plank exercise were less sanguine about their efficacy at the task than those who began with the wall-sit exercise.

Figure 1a. Results from Messé et al. (2000) Experiment 1.

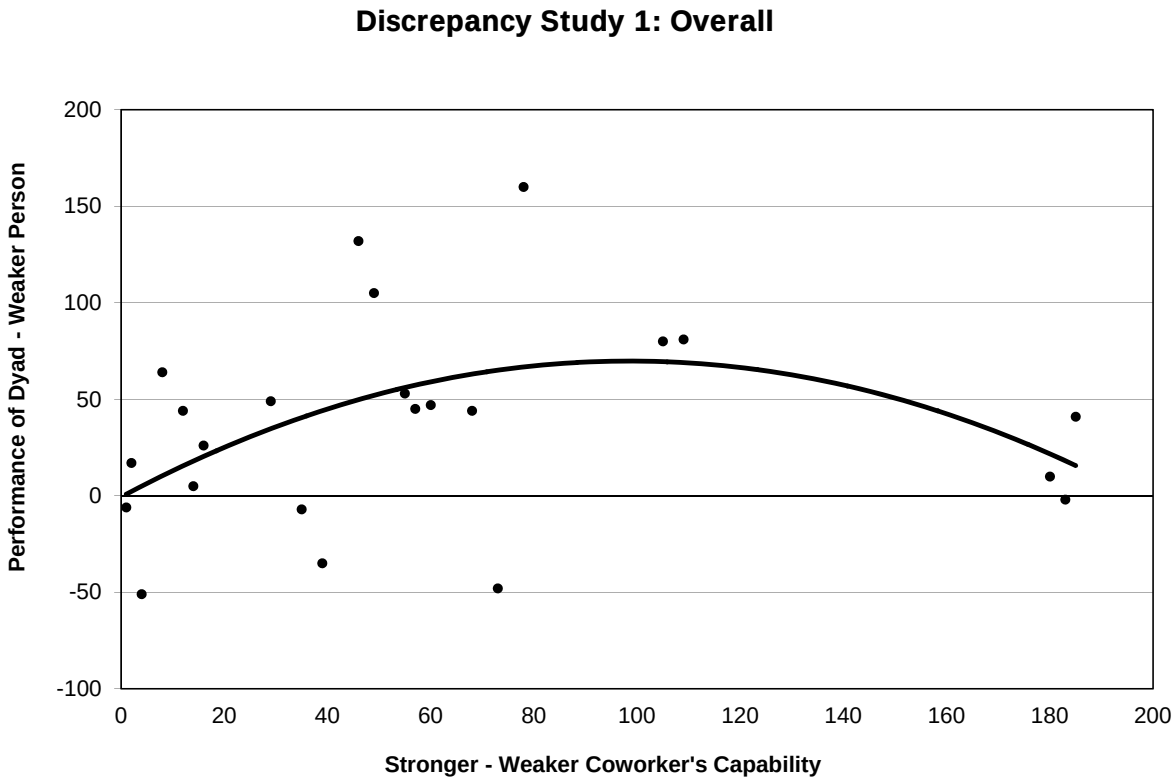

Figure 1b. Results from Messé et al. (2000) Experiment 2.

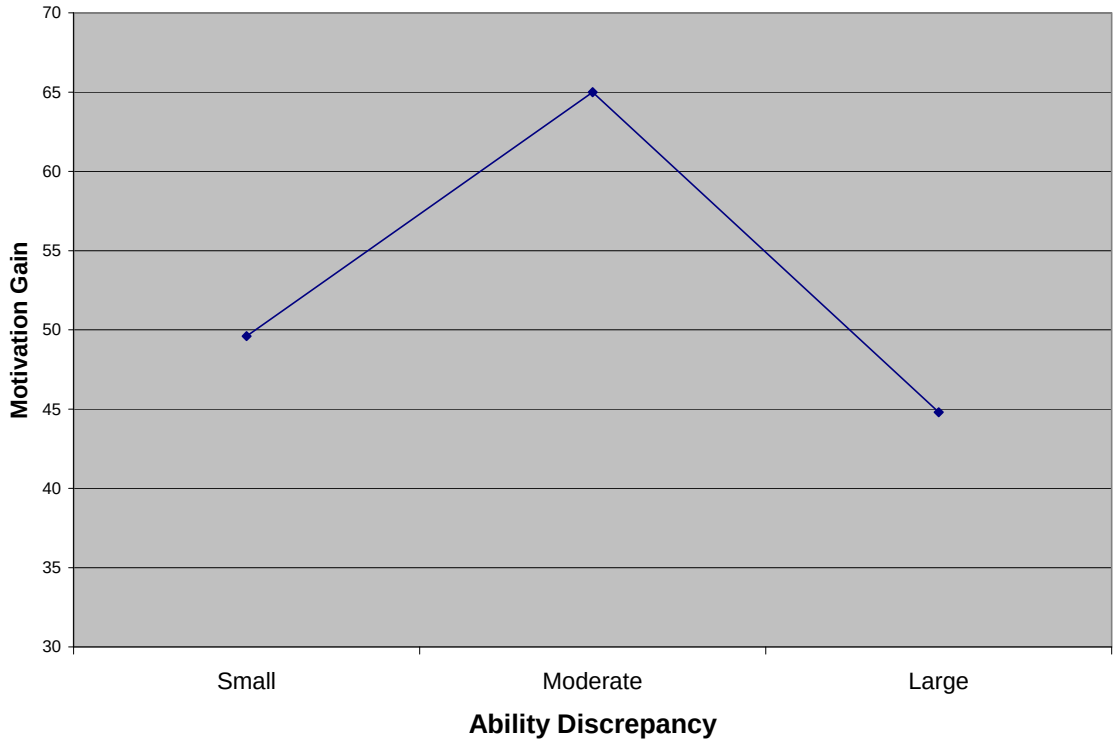

Figure 2. Results from Kerr et al. (2007).

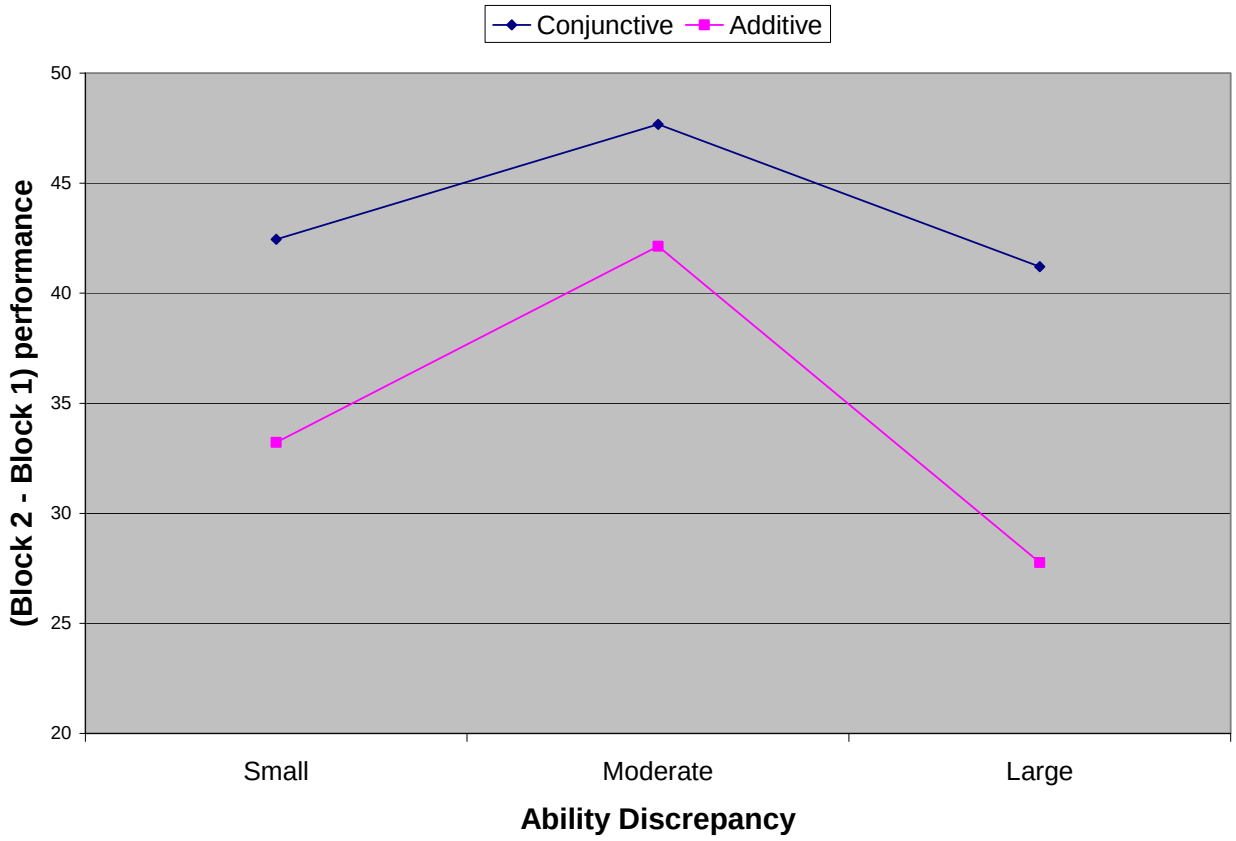

Figure 3. Contrast of Individuals vs. S-always-inferior conditions across blocks.

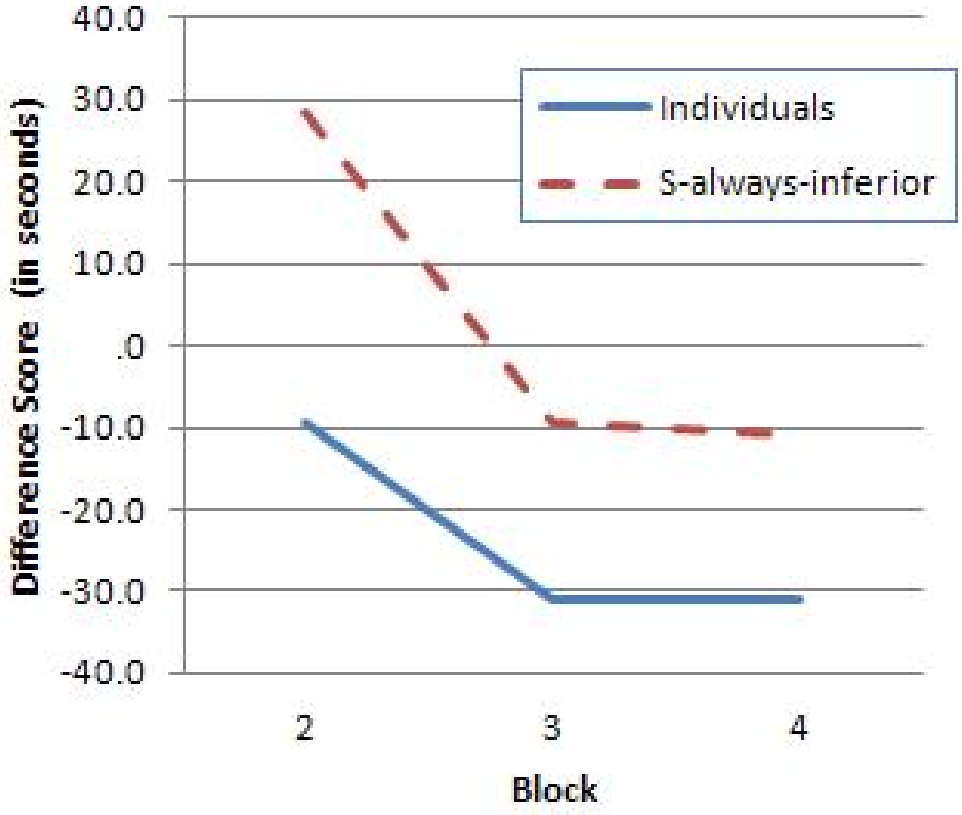

Figure 4a. Contrast of S-always-inferior vs. S-superior-on-wall-sit conditions, Plank exercise performance.

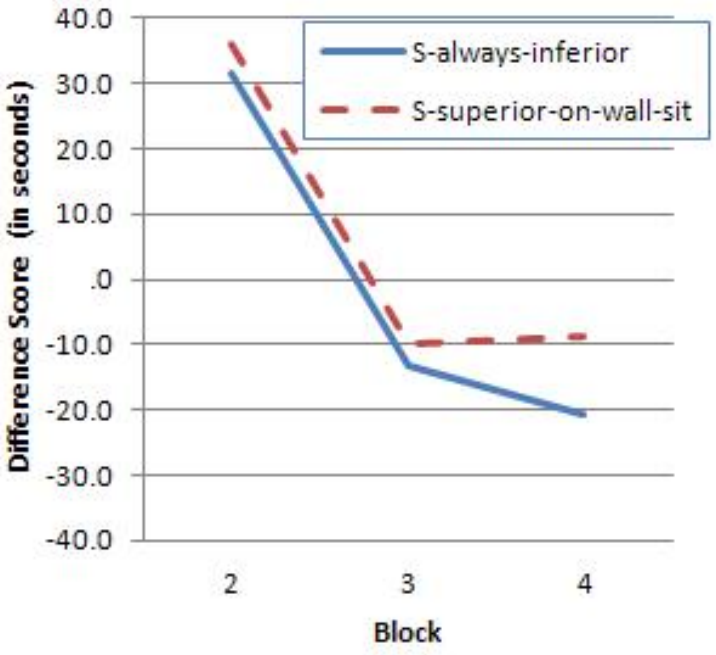

Figure 4b. Contrast of S-always-inferior vs. S-superior-on-plank conditions, Wall-sit exercise performance.