Peer Tutoring in a Physical Education Setting: Influence of Tutor Skill Level on Novice Learners’ Motivation and Performance

Fabienne d’Arripe-Longueville, Christophe Gernigon, Marie-Laure Huet
National Institute of Sport and Physical Education, Paris

Marielle Cadopi  Fayda Winnykamen
University of Montpellier I  University of Paris V

Based on Vygotsky’s theory of cognitive development and its concept of zone of proximal development, this study examined how the skill level of a peer tutor affects the achievement motivation of novice learners and their performance in a swimming task. Gender differences were also explored. High school students (N = 48) were assigned in a 2 x 3 (Gender x Tutor skill level: novice vs. intermediate vs. skilled) factorial design. Participants were invited to observe a same-sex peer tutor, complete a self-efficacy questionnaire, train with their tutor for 8 minutes, and complete a goal involvement questionnaire. Results demonstrated that skilled tutors yielded the best swimming skills for boys, whereas skilled and intermediate tutors yielded better skills than did novice tutors for girls. The skilled tutor group led to higher self-efficacy for improvement and gave more demonstrations and verbal information than did the novice group. Male tutees adopted higher ego involvement goals and trained more physically, whereas female tutees adopted higher learning goals and received more demonstrations and verbal instructions. Results are discussed in relation to educational studies conducted in a Vygotskian perspective.

Key Words: peer-assisted learning, observational learning, self-efficacy, gender

Collaborative learning, peer tutoring, and peer modeling are well recognized as successful strategies through peer interaction in educational settings (Topping & Ehly, 1998). Such peer-assisted learning methods refer to a set of alternative teaching arrangements in which students serve as instructional assistants for classmates or other children (Maheady, Harper, & Mallette, 1991). The cost-effectiveness of these procedures and the contemporary importance of student-centered learning have led to increased research interest in peer-assisted learning in various school...
settings. Specifically, studies of collaborative learning in physical education contexts have been shown to be innovative interventions that maximize classroom resources and promote a more effective and more dynamic learning experience (e.g., Dyson, 2001; Johnson & Ward, 2001; Siedentop & Tannehill, 2000).

Peer tutoring, the most widely known peer-assisted learning method, is characterized by specific role-taking: one peer has the role of tutor whereas the other has the role of tutee (Topping & Ehly, 1998). Research into peer tutoring set within Vygotsky’s theory (1962, 1978) considers that children are introduced to new patterns of thought and new understandings by engaging in dialogue with others. Vygotsky emphasized asymmetrical relationships in which one interactant is more knowledgeable and expert than the other. He postulated the existence of a **zone of proximal development**, which is the zone bounded by what the child is capable of doing alone and by the level achieved when solving problems in collaboration with someone who is more experienced.

Although Vygotsky referred more to adults and focused more upon culture as providing tools for thinking (Gillen, 2000), educational studies set within the Vygotskian framework have indicated that collaboration with peers might have a positive effect on children’s cognitive development (Hogan & Tutge, 1999; Schacter, 2000). Classwide settings (Fuchs, Fuchs, Mathes, & Simmons, 1997) and one-on-one settings using both same-age (Cohen, Kulik, & Kulik, 1982) and cross-age (Osguthorpe & Scruggs, 1986) peer tutoring have demonstrated their significant academic benefits for both helpers and helped. Some reasons put forward for the positive academic effects of tutoring are that it enhances self-concept and feelings of self-worth (Fantuzzo, King, & Heller, 1992) and makes learning more fun and exciting (Maheady, 1998).

The benefits of peer tutoring depend on a complex set of factors that include the age and ability level of the children and of their partner, the children’s motivation to collaborate, the nature of the task, and the institutional and cultural support for collaboration (Hogan & Tutge, 1999). An important factor is the competency levels between peers. Based on Vygotsky’s theory (1962, 1978), studies have indicated that children who work with a more competent partner generally improve more than those who have an equally or less competent partner (see Hogan & Tutge, 1999, for a review). The concept of zone of proximal development suggests that neither the task difficulty nor the guidance given to the children be too far in advance of their current level of ability. For instance, Tutge, Winterhoff, and Hogan (1996) tailored the problems given to target children in each dyad so that the most difficult problem could be solved by using the rule one higher than they had used at pretest. Their partners had used a rule no more than two higher than that of the target children. This kind of pairing was more effective in facilitating cognitive growth than had been the case in previous studies when the problems were not tailored in this way (Tutge, 1989, 1992).

Gender is generally underexplored as a potential factor influencing peer tutoring. A few studies of cooperative learning in the academic domain found that, compared to boys, girls have greater achievement and increased perceived status in classrooms with cooperative goal structures (e.g., Johnson, Johnson, & Stanne, 1985). Girls also express a greater preference for cooperative learning instruction at school, whereas boys are more attracted by competitive and individualistic learning conditions (e.g., Owens & Barnes, 1982).

These gender preferences have been identified regarding social interaction
modes. The literature in educational psychology (e.g., Charlesworth & Dzur, 1987) indicates that in dyadic situations, girls exhibit more sensitivity to their partner than do boys and prefer to engage in teaching/learning activities (tutoring) or sharing (cooperation). Boys, on the other hand, are more inclined to direct their energy toward acquiring knowledge individually or through social comparison. In a study set within a Vygotskian framework, Tutge (1992) reported that girls were more likely to regress in a balance beam task, following collaboration with other girls, than were boys who had worked with boys, in part because the girls seemed more interested in preserving good relationships with their partners than in arguing with one another. The findings of these studies indicate that gender is an area deserving greater attention.

Although considerable research has been carried out by developmental and educational psychologists on peer tutoring in academic contexts, very few empirical studies have been conducted in physical education. The majority of studies have been conducted in adapted physical education or within inclusive settings pairing typically developing peers with students who have developmental disabilities (e.g., Houston-Wilson, Dunn, van der Mars, & McCubbin, 1997; Webster, 1987). Because the advantages of peer tutoring may be linked to the context in which they occur, several recent studies have examined the appropriateness of peer tutoring as an effective strategy for typically developing peers in inclusive and intact class settings. Specifically, studies using an independent variable called peer-mediated accountability (PMA) have reported the effects of training students to assess each other’s performance in physical education (Crouch, Ward, & Patrick, 1997; Ward, Smith, Makasci, & Crouch, 1998). These studies suggested that PMA is a good strategy for developing opportunities to practice newly acquired skills, while classwide peer tutoring might be a better instructional strategy if students are learning new skills (Johnson & Ward, 2001). In light of previous research in the educational psychology literature, there is an obvious need to extend our knowledge about the factors that would determine the effectiveness of pairing students together in physical education settings.

The role of the skill level of a peer in the sport domain has been examined through the sociocognitive theory of observational learning (Bandura, 1986) and the self-efficacy theory (Bandura, 1977, 1997), and discrepant findings have been produced (see McCullagh & Weiss, 2001, for a review). Viewing a skilled model has been shown to entail higher self-efficacy and better motor performance (Lirgg & Feltz, 1991) and a more exact pictorial representation of the skill (d’Arripe-Longueville, Fleurance, & Winnykamen, 1995) than viewing an unskilled model. An additional role of the model skill level could be that it may be viewed
as a motivational cue used by observers to inform them about their own ability to achieve a task. According to Yando, Seitz, and Zigler (1978), the ideal condition of modeling could be when the task that is being demonstrated appears to involve a skill level that is just beyond the observer's actual level of competence. Based on this notion of optimal challenge, consistent with the aforementioned Vygotskian concept of zone of proximal development, d’Arripe-Longueville et al. (1995) examined the influence of competency levels between peers when learning a forward somersault in dyads. They found that a weak competence asymmetry between peers entailed a better pictorial representation of a motor skill than a strong asymmetry, while there appeared to be no difference between weak and strong asymmetry on performance outcomes.

To extend these findings, it would be interesting to examine the influence of peer skill level in physical activity contexts in relation to situational indexes of achievement motivation and performance. Specifically, since previous studies in educational psychology have indicated that peer tutoring has a favorable effect on self-concept (Fantuzzo et al., 1992; Topping & Ehly, 1998), self-efficacy would be an important construct to consider. Furthermore, research that has dealt with the predictions of self-efficacy and those of achievement goal theory (Bandura, 1997; Schunk & Swartz, 1993) also suggests that states of goal involvement—the goals that individuals pursue in a specific situation and that refer to self-referenced or norm-referenced criteria (Nicholls, 1989)—should also be evaluated.

Therefore, the main purpose of the present study was to explore whether there is an optimum level of challenge between peers that will maximize motivated behavior and performance in a sport setting. Specifically, we examined how the skill level of a peer tutor affects the novice learner’s self-efficacy, states of goal involvement, social interactions, and performance in a swimming task. Weak skill-level asymmetry (interaction with an intermediate-skilled peer tutor) was hypothesized as being most likely to create an optimal learning challenge compared to skill-level symmetry (interaction with a low-skilled tutor), or strong skill-level asymmetry (interaction with a skilled peer tutor). Thus intermediate peer tutors were expected to entail the highest scores in self-efficacy, learning goals, relevant guidance, and performance.

A secondary aim of this study was to explore potential interactions between gender and tutor skill level. Because swimming is considered a gender neutral activity (Koivula, 1995), no differences were expected between genders in feelings of self-efficacy (Lirgg, 1991). However, according to the literature showing that females are generally more task-oriented and less ego-oriented than males (Duda, 1988; White & Duda, 1994), it was hypothesized that females working with intermediate tutors would be the most task-involved while males were expected to be the most ego-involved. Consistent with previous research in the educational psychology literature (Charlesworth & Dzur, 1987; Tutge, 1992), it was also expected that females paired with intermediate peers would benefit from the most verbal information whereas males working with intermediate peers would engage the most in physical behavior. There was not enough information in the literature to hypothesize about how the experience of peer tutoring might result in different performance outcomes for male and female students.

Finally, correlations between relevant variables pertaining to Bandura’s (1997) self-efficacy theory were examined. Specifically, self-efficacy beliefs were expected to be positively related to swimming performance and to task-involvement goals.
Method

Participants and Design

The participants were 48 French students, 24 males and 24 females. Their average age was 18.3 years (SD = 1.2). The students were predominantly Caucasian and most were from working and middle-class socioeconomic backgrounds. They were selected after a pretest involving 306 twelfth-graders who attended three high schools in Paris. The students volunteered to participate in the study, and the parents gave written consent for them to do so. The participants were told that they would be videotaped for the purposes of this experiment but that their confidentiality would be protected.

Novices in the swimming skill to be learned, the breaststroke turn, were chosen because paying attention to instructions and demonstrations is of particular interest during the early phases of acquiring a motor skill (Fitts & Posner, 1967; Schmidt, 1988). Swimmers were videotaped performing two turns and were evaluated independently by the judges, who were blind to the purpose of the study and experimental conditions. The participants were those for whom two expert judges in swimming fully agreed to give for both turns a score of 1 on a 4-point scale, indicating a non-hydrodynamic body position without propulsion. The 48 participants were assigned to the six independent groups of a 2 x 3 design (Gender x Tutor skill level: novice vs. intermediate vs. skilled). Participants were not informed of their skill level nor that of the peer they were paired with.

Peer Tutors

Forty eight students (24 M, 24 F, mean age = 18.5 years, SD = 1.4) served as peer tutors. Tutors and tutees were part of the same classrooms. Same-sex familiar peers were chosen in order to maximize perceived similarity between peers during the peer modeling phases (Bussey & Bandura, 1984). Their swimming skills were assessed in a pretest following the aforementioned procedure. Sixteen students who obtained a form score of 1 on the 4-point scale were identified as novice tutors. Sixteen who obtained a score of 2 served as intermediate tutors, and 16 students who obtained a form score of 3 were identified as skilled tutors.

Study Setting and Task Materials

The study took place in regularly scheduled physical education classes. The task to be learned was a breaststroke turn in swimming. It was chosen because it could be easily modeled and was part of the participants’ curriculum. The task was to be learned in a swimming lane of a 25-m pool in which the minimal depth was 1.80 m. All students in each class were in the pool at the same time. However, only the selected dyads for the study came into the experimental lane, one after the other, while the others were taught by their regular physical education teacher in the other part of the pool. The instructions for the participants were written on a white board placed 5 m from the side of the pool. The board also served as the starting point for the participants.

A Sony Hi8 portable video camera was used to film each participant’s per-
formance. To ensure accuracy, the camera operator was placed 3.00 m from the wall where the swimmers were to turn, and 1.50 m from the side of the pool. Each swimmer was videotaped from a profile position during the pretest and the posttests. An experimenter gave instructions. Traveling shots were used to film the behaviors of each member of a dyad during the training session. Another experimenter used a directional microphone and took notes to record the verbal exchanges within each dyad.

**Measures**

Measures included swimming skill, self-efficacy beliefs, and states of goal involvement. According to a previous qualitative investigation of peer interaction in a similar swimming setting with same-age participants (d’Arripe-Longueville, Gernigon, Huet, Winnykamen, & Cadopi, 2002), the number of task attempts, number of demonstrations, and verbal information during the training phase were also measured. For all these measures, interobserver agreements were calculated and are presented in the results.

**Swimming Skill.** A form score was calculated for each breaststroke turn carried out at all three assessment periods (pre-, post-, and retention test), using the independent ratings of two physical education teachers who were expert swimmers and who were blind to the study purpose and experimental conditions. The raters had participated in a pilot study in which 80 high school seniors carried out the same breaststroke turn individually. From this study, four stages of acquisition of this skill were identified. This resulted in a 4-point scale of 8 half-point increments. Each score on the scale was defined by drawings and texts and had been validated by three other expert swimmers.

A score of 1 indicated that the wall was touched with one hand and that the swimmer glided through the water with a non-hydrodynamic body position, without using arm or leg movements to propel him or her. A score of 2 indicated that the wall was touched with both hands and the swimmer glided through the water by doing a breaststroke cycle. A score of 3 indicated that the turn was carried out correctly (i.e., touching the wall of the pool with both hands, gliding through the water by pulling both arms back strongly while at the same time doing a scissors kick to propel the swimmer out of the water). A score of 4 corresponded to an elite level in that the turn was carried out correctly and with efficacy.

The raters practiced the scoring scheme using pilot videotapes of 15 same-age students. The participants were coded independently and, because the raters were very familiar with the scale, a 100% agreement was obtained. For subsequent data analysis, the videotapes were viewed separately. The judges averaged the scores of the two turns for each participant and interrater reliability was calculated.

**Self-Assessed Swimming Skill.** The self-assessed swimming skill inventory consisted of the same aforementioned scale, representing the four levels of technique of the breaststroke turn. Participants were asked to mark the average score they thought they had reached after carrying out the breaststroke turn twice.

**Swimming Self-Efficacy Beliefs.** A swimming self-efficacy scale was created to measure participants’ efficacy to perform the breaststroke turn. Level and strength of efficacy judgments were measured by adapting procedures of previous research (e.g., Bandura & Adams, 1977). The level scale represented the four levels of technique of the breaststroke turn. Participants were asked to mark the best score
they expected to get in the breaststroke turn after the training session. Although single-item measures are less reliable than multiple-item measures, this kind of measure was necessary because in the skill to be learned, reaching one level does not inevitably imply succeeding at the lower ones. Furthermore, although participants did not receive any training on the self-efficacy measure, they had to assess their initial form scores after the pretest on a similar scale. This was supposed to help them determine their potential scores, since the same components made up the form scores and the related self-efficacy scores in both scales.

To measure the strength of self-efficacy, the participants indicated their degree of confidence about reaching the expected score on a scale ranging from 10% (not sure) to 100% (totally sure). Self-efficacy for improvement (Schunk & Hanson, 1985) was also indirectly assessed by using the self-assessed swimming skill scores as a covariate.

**States of Goal Involvement.** States of achievement goal involvement were measured with two subscales, namely learning and ego involvement goals, of Thill and Crevoisier’s (1994) Personal Standards Evaluation Questionnaire (PSEQ). The PSEQ in its original form consists of 18 items scored on a Likert scale from 1 = strongly disagree to 7 = strongly agree, and distinguishes self-referenced standards and socially-referenced standards. Instructions and items in this questionnaire prompt the participants to focus on the goals they adopted when achieving a particular task. This questionnaire has indicated satisfactory reliability with alpha coefficients varying between 0.82 and 0.85 for all scales and satisfactory construct validity (see Thill & Cury, 2000, for details of the psychometric properties of this questionnaire).

As our purpose was to verify that the instructions for improvement actually elicited the corresponding goals in students, we retained only the learning goals and the ego involvement goals subscales, similar to Nicholls’ (1984, 1989) task and ego involvement conceptualizations. The items related to the learning goals address learning or personal progress (e.g., I carried out this exercise to improve on my weaknesses and my level). The items related to ego involvement goals focus on judgment from others, displaying one’s ability or showing greater ability than others (e.g., I participated in this exercise mostly to show others that I am the best).

**Number of Attempts.** Recent research in the sport domain has used the number of attempts as an indicator of physical engagement (e.g., Gernigon, Thill, & Fleurance, 1999; Solmon, 1996). Participants’ physical engagement has been considered here to be the number of attempts, both partial (glide and return to swimming) and complete (approach/glide/return to swimming), made during the 8 minutes allocated for practicing the turn. The number of attempts was counted independently by two researchers who were former physical education teachers and experienced in qualitative methods. The raters were trained to code for behaviors using pilot videotapes.

**Demonstrations.** All demonstrations the tutor provided to the participant during the 8-minute training session were counted, whether complete or partial, accompanied or not accompanied by verbal explanations (d’Arripe-Longueville, 1998; d’Arripe-Longueville et al., 2002). The frequency of demonstrations was calculated using the same procedure as for the number of attempts.

**Verbal Information.** The frequency of verbal information given to the participant was counted. This verbal information included (a) clarifying the task goal,
(b) verbally describing the required movements or level of achievement, and (c) expressing a value judgment (d’Arripe-Longueville et al., 2002).

**Procedure**

The pretest consisted of carrying out a breaststroke turn twice. The students were tested individually and were given directions as follows:

You are going to do two breaststroke turns in accordance with the following instructions: (1) begin at the 5.00-m mark and breaststroke to the side of the pool; (2) touch the side with both hands before turning; (3) glide through the water by pulling your two arms back as strongly as possible while at the same time doing a scissors kick propelling you out of the water; (4) do two breaststrokes at the surface; (5) rest, and tell me when you are ready to do your second attempt.

The experimenter invited each student to reread the instructions as written on the board. The students were given time to do so and to ask any questions they had concerning the task so as to be sure they understood it completely. Each student was then asked to carry out the task twice consecutively, beginning whenever he or she was ready. It was felt that two trials were needed to check the students’ form scores. These trials were videotaped to allow subsequent objective rating of skills and calculation of rater reliabilities.

After the pretest, each participant had to assess his or her own swimming skill (form scores) on a 4-point scale that described and illustrated the different stages of learning the breaststroke turn. The participants were then asked to observe their partner peer perform the turn twice, and to fill in a self-efficacy questionnaire concerning the form score they expected to reach after training with their partner. The training session consisted of training freely with the tutor for 8 minutes. Pilot studies indicated this amount of time was best suited to observing progress, given the respiratory constraints of the task which quickly brought on fatigue. The tutors’ interventions were not structured. The instructions were as follows:

You will now work with your classmate for 8 minutes to improve your turn as much as possible. You are free to organize your work in whatever way you wish. We will record what you do, but the document will remain confidential. Is everything clear? Do you have any questions?

After the training session, participants were asked to fill in two subscales of the Personal Standards Evaluation Questionnaire (Thill & Crevoisier, 1994; Thill & Cury, 2000), which allows goal involvement to be evaluated. Finally, participants performed two breaststroke turns under the same conditions as for the pretest immediately following the training session (posttest) and again 2 weeks later (retention test). They did not practice any swimming during the 2-week interval between posttests.

**Data Analyses**

Interobserver agreement (IOA) for the frequency of attempts, demonstrations, and verbal information was conducted on all observations and was calculated using the following formula: $$IOA = \frac{agreements}{(agreements + disagreements)}$$
100 (Thomas & Nelson, 1996, p. 228). Interrater reliability for performance form judgments was assessed by calculating the percentage of agreement between the two judges and the intraclass correlation coefficient ($R$) using a two-way ANOVA (Participants x Raters) with repeated measure on the last factor (Thomas & Nelson, p. 221).

For the changes in swimming skill, a $2 \times 3 \times 3$ (Gender x Tutor skill level x Assessment period: pretest vs. posttest vs. retention test) repeated-measures analysis of variance (RM ANOVA) was conducted. In order to control for inflated alpha, Bonferroni’s corrections were used, and the adjusted alpha necessary to demonstrate significance was $p < .0083$ (Thomas & Nelson, p. 173). For self-efficacy for improvement, a $2 \times 3$ (Gender x Tutor skill level) analysis of covariance (ANCOVA) that used self-assessed swimming skill as the covariate was processed. Because self-efficacy and physical engagement were conceptually linked (Bandura, 1997), a $2 \times 3$ multivariate analysis of variance (MANOVA) was conducted to examine gender and tutor skill level differences.

Goal involvement variables (learning and ego goals) and tutor behaviors (verbal information and demonstrations) were examined in two other MANOVAs. Discriminant function analyses (Huberty & Wisenbaker, 1992) were then used to identify which of the dependent variables maximized tutor skill level and gender differences. Effect sizes ($ES$) were also calculated using pooled standard deviations (Hedges & Olkin, 1985). Pearson correlational analyses were computed among theoretically relevant variables.

Results

Preliminary Reliability Analyses

In the present study, the internal consistency of the two subscales of the PSEQ (Thill & Crevoisier, 1994), as assessed with Cronbach’s alpha, were .75 for the learning goals subscale and .81 for the ego-involvement goals subscale.

For interrater reliability, analyses yielded a high percentage of agreement at both posttest (92.18%) and retention test (93.75%), and good intraclass correlations ($R = .935$ and .946, respectively). Therefore the judges’ ratings of performance form were averaged. Perfect agreement was obtained for the number of attempts of carrying out the task. For the frequency of demonstrations, the reliability score was 97.2%. For the frequency of verbal information, it was 92.8%.

Tutor Skill Level, Gender, and Assessment Period Differences

Table 1 lists the means and standard deviations of all measures for each experimental condition.

Swimming Skill. On the measure of swimming skill, the RM ANOVA indicated main effects for assessment period, $F(2, 78) = 188.40, p < .001$; for tutor skill level, $F(2, 78) = 15.78, p < .001$; and for gender, $F(1, 78) = 9.18, p < .005$. The analyses also yielded a Gender x Tutor skill level interaction, $F(2, 78) = 7.75, p < .001$; a Gender x Assessment period interaction, $F(2, 78) = 8.48, p < .001$; and a Tutor skill level x Assessment period interaction, $F(4, 78) = 11.51, p < .001$. Follow-up paired $t$-tests ($p < .05$) indicated that for assessment period effects, all groups improved in swimming skill from pretest to posttest, and from pretest to retention test. However, no differences were found between posttest and retention test.
Table 1  Means (± SD) for Dependent Measures (N = 48)

<table>
<thead>
<tr>
<th>Dependent measure</th>
<th>Novice peer tutor (Skill symmetry)</th>
<th>Intermediate peer tutor (Weak skill asymmetry)</th>
<th>Skilled peer tutor (Strong skill asymmetry)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
<td>Males</td>
</tr>
<tr>
<td>Intermediate swimming skill</td>
<td>1.43 (.41)</td>
<td>1.43 (.41)</td>
<td>1.81 (.25)</td>
</tr>
<tr>
<td>Delayed swimming skill</td>
<td>1.56 (.41)</td>
<td>1.43 (.41)</td>
<td>2.06 (.32)</td>
</tr>
<tr>
<td>Self-assessed swimming skill</td>
<td>1.50 (.54)</td>
<td>2.00 (.63)</td>
<td>2.00 (1.00)</td>
</tr>
<tr>
<td>Self-efficacy level</td>
<td>2.50 (.53)</td>
<td>2.62 (.91)</td>
<td>2.37 (.51)</td>
</tr>
<tr>
<td>Self-efficacy strength</td>
<td>52.50 (12.81)</td>
<td>47.50 (13.88)</td>
<td>50.00 (13.09)</td>
</tr>
<tr>
<td>Learning goals</td>
<td>5.07 (1.76)</td>
<td>4.62 (.95)</td>
<td>4.73 (.99)</td>
</tr>
<tr>
<td>Ego goals</td>
<td>3.85 (1.57)</td>
<td>2.81 (.85)</td>
<td>3.10 (1.74)</td>
</tr>
<tr>
<td>Number of attempts</td>
<td>11.87 (6.7)</td>
<td>7.25 (1.83)</td>
<td>9.62 (4.43)</td>
</tr>
<tr>
<td>Demonstrations</td>
<td>0.37 (.21)</td>
<td>1.00 (.92)</td>
<td>1.32 (.21)</td>
</tr>
<tr>
<td>Verbal information</td>
<td>2.12 (1.95)</td>
<td>8.00 (6.61)</td>
<td>3.90 (1.11)</td>
</tr>
</tbody>
</table>

*Note: Swimming skill, self-assessed swimming skill, and level of self-efficacy were rated on a 1-to-4 scale. Strength of self-efficacy scores ranged from 1 to 100. Self-efficacy for improvement scores ranged from 0 to 3. States of goal involvement were rated on a 1-to-7 scale.*
Effect sizes computed within each group from pretest to posttest and from pretest to retention test indicated large $ES$ for all groups. Specifically, the skilled tutor group showed larger $ES$ ($ES = 3.86$ and $4.64$) than both the intermediate group ($ES = 3.47$ and $3.20$) and the novice group ($ES = 1.48$ and $1.69$). With regard to the tutor skill-level main effect, results indicated that compared to learners paired with novice tutors, those paired with skilled tutors performed significantly better at posttest and retention test ($ES = 1.60$ and $1.14$, respectively). The skilled-tutor group also performed better than the intermediate-tutor group at posttest and retention test ($ES = .76$ and $.72$, respectively).

With regard to gender, boys scored higher than girls both at posttest ($ES = .43$) and retention test ($ES = .86$). Additionally, both at posttest and retention test, boys in the skilled-tutor condition demonstrated significantly better skills than their female counterparts (all $p < .001$, $ES = 1.82$ and $2.77$, respectively) and better than boys assigned to the intermediate-tutor (all $p < .01$, $ES = 1.82$ and $1.82$) or novice-tutor (all $p < .001$, $ES = 2.44$ and $2.12$) conditions. At retention test, both genders in the intermediate-tutor condition performed respectively higher ($p < .05$, $ES = 1.36$ and $.97$) than their counterparts in the novice tutor condition.

**Self-Efficacy Beliefs and Physical Engagement.** A $2 \times 3$ MANOVA was conducted to examine gender and tutor skill level differences in self-efficacy and physical engagement (i.e., number of task attempts). Results indicated only a significant main effect for gender, Wilks’ $\Lambda = .783$, $F(3, 40) = 3.71$, $p = .01$. The standardized discriminant coefficients indicated that the number of attempts (.57) maximized the gender differences, while the level (.27) and the strength (.25) of self-efficacy for form scores did not contribute significantly to these differences. An examination of mean differences revealed that boys displayed more task attempts than girls ($M = 11.04$ vs. $7.50$; $ES = .95$).

With regard to self-efficacy for improvement, the ANCOVA yielded a significant effect for the covariate, $F(1, 41) = 10.85$, $p < .01$, and a main effect for tutor skill level, $F(2, 41) = 4.43$, $p < .05$. Follow-up one-way ANCOVAs indicated that learners assigned to the skilled-tutor condition reported higher self-efficacy for improvement than those assigned to the novice condition ($p < .05$, $ES = 1.81$).

**States of Goal Involvement.** Results of the $2 \times 3$ ($Gender \times Tutor$ skill level) MANOVA indicated only a significant main effect for gender, Wilks’ $\Lambda = .765$, $F(2, 41) = 6.28$, $p < .01$. The standardized discriminant coefficients indicated that both learning goals (.63) and ego goals (.43) maximized gender differences. An examination of mean differences revealed that girls adopted higher learning goals than boys ($M = 5.23$ vs. $4.61$; $ES = .55$), whereas boys adopted higher ego goals than girls ($M = 3.7$ vs. $2.61$; $ES = .82$).

**Verbal Information and Demonstrations.** Results of the $2 \times 3$ ($Gender \times Tutor$ skill level) MANOVA indicated a significant main effect for gender, Wilks’ $\Lambda = .765$, $F(2, 41) = 6.28$, $p < .01$, and a significant main effect for tutor skill level, Wilks’ $\Lambda = .765$, $F(2, 41) = 6.28$, $p < .01$. However, the $Gender \times Tutor$ skill level interaction effect was not significant, Wilks’ $\Lambda = .971$, $F(2, 25) = .18$, $p < .94$. A follow-up discriminant function analysis revealed that verbal information (.55) and demonstrations (.50) both maximized gender differences. An examination of mean differences indicated that girls received more verbal information than boys ($M = 11.8$ vs. $3.38$; $ES = 1.46$). Demonstrations were also found to be more frequent among girls than among boys ($M = 3.04$ vs. $1.93$; $ES = .73$). Furthermore, discriminant
function analysis indicated that verbal information (.66) and demonstrations (.55) were significant contributors to tutor skill level differences.

Post hoc Scheffé tests also revealed that participants in the novice-tutor group were given significantly less verbal information than those in the intermediate-tutor group ($M = 5.06$ vs. $7.85$, $p < .05$) and those in the skilled-tutor group ($M = 8.81$, $p < .01$). These effects were large ($ES = .71$ and $1.21$, respectively). Skilled tutors ($M = 2.62$) and intermediate tutors ($M = 2.72$) also provided more demonstrations than novice tutors ($M = .68$, both $p < .01$). Effect sizes were large ($ES = .93$ and $1.21$, respectively).

**Correlational Analyses**

The level of self-efficacy was positively related to immediate and delayed swimming skills ($r = .30$, $p < .05$, and $r = .39$, $p < .01$, respectively). The level and strength of self-efficacy were both correlated with the number of attempts ($r = .30$ and $.35$, respectively, both $p < .05$). Self-efficacy for improvement also correlated positively with immediate and delayed swimming skills ($r = .34$ and $.30$, respectively, both $p < .05$). Surprisingly, the learning goals correlated negatively with level of self-efficacy ($r = −.35$, $p < .01$), number of attempts ($r = −.60$, $p < .01$), and immediate and delayed performance (both $r = −.29$, $p < .01$). Furthermore, ego involvement goals correlated positively with strength of self-efficacy ($r = .42$, $p < .01$).

Verbal information correlated positively with number of demonstrations ($r = .70$, $p < .001$), but negatively with number of attempts ($r = −.40$, $p < .05$). Finally, number of attempts correlated positively with delayed swimming skills ($r = .36$, $p < .05$). Correlations were also computed among genders separately. For boys, ego involvement correlated positively with strength of self-efficacy ($r = .60$, $p < .01$), while learning goals were negatively related to level of self-efficacy ($r = −.48$, $p < .05$), number of attempts ($r = −.70$, $p < .001$), and delayed swimming skills ($r = −.56$, $p < .01$). For girls, no significant relationships were found between goals and self-efficacy, but learning goals correlated positively with immediate and delayed swimming skills ($r = .42$ and $.39$, respectively, both $p < .05$).

**Discussion**

The primary purpose of the present study was to explore whether there is an optimum level of challenge between peers that would maximize motivated behavior and performance in a physical education setting. Specifically, the influence of a peer tutor’s skill level on a novice learner’s situational indexes of achievement motivation and swimming performance was examined. Further, the study was designed to explore potential interactions between gender and tutor skill level. The findings revealed only partial support for our predictions.

Significant results demonstrated that skilled tutors yielded the best swimming skills for boys, whereas skilled and intermediate tutors both yielded better skills for girls than did novice tutors. The skilled-tutor group showed higher self-efficacy for improvement, more demonstrations and verbal information, and higher performance than the novice group. Skilled tutors also led to higher performance than did intermediate tutors. Boys adopted higher ego-involvement goals, trained more physically, and performed better than girls, who in turn adopted higher learning
goals and received more demonstrations and verbal instructions.

In accord with Vygotsky’s framework (1962, 1978) and previous studies in cognitive learning (Tudge et al., 1996) and in the sport domain (d’Arripe-Longueville et al., 1995), we had hypothesized that weak skill-level asymmetry between peers (interaction with an intermediate-skilled peer) would create the best conditions for help and thus yield higher motivation and performance than would skill-level symmetry (novice peer tutor) or strong skill-level asymmetry (skilled peer tutor). As we expected, learners training with intermediate-skilled peers performed better than those training with novice peers. However, the skilled-tutor group led to the most adaptive patterns, especially for boys, and also showed the strongest effect sizes. These findings are consistent with the conclusions of earlier studies in the educational psychology literature which have indicated that children working with a more competent partner improve more than those who have an equally competent partner (Tutge, 1992; Tutge et al., 1996).

These results also support previous findings in the sport psychology literature on modeling which have shown that a correct model leads to higher self-efficacy and performance than does an unskilled model (Lirgg & Feltz, 1991), and that gains in competence are greater in asymmetric-competence conditions than in symmetric ones (d’Arripe-Longueville et al., 1995). Also, the fact that participants working with skilled or intermediate peers received more help through demonstrations and verbal information than those working with novice peers suggests that due to their lesser skill level, the tutors in the symmetrical conditions were not as good at providing guidance as the more skilled peers. This hypothesis supports the idea that guidance techniques develop along with the skill level of the participants, as has been shown in predominantly cognitive tasks (Frayss, 1985).

Contrary to our predictions, weak skill-level asymmetry (intermediate peer tutors) did not lead to higher motivation and performance than strong asymmetry. Conversely, boys paired with skilled tutors outperformed boys in the intermediate group, while no differences were found between these conditions for girls. There are several possible reasons for why the expected tutor skill-level differences did not emerge. Perhaps the relatively arbitrary strong asymmetry design of the present study was not perceived as such by the students, and may have provided the optimal learning challenge that was expected in the intermediate condition. This would mean that the concept of weak skill-level asymmetry should be tailored to each skill to be acquired, and that the perceived competence gap between peers warrants assessment in future work.

A related explanation might be that the motor skill to be learned and the design of the present study were not totally appropriate to a constructivist approach. In fact, learning the breaststroke turn in this study consisted more in form reproductions than in problem-solving as defined in the developmental psychology literature. Because our design emphasized modeling, it can be assumed (Schmidt, 1988; Sheffield, 1961) that, for novice learners, the more exact symbolic representation provided by skilled peer tutors was more likely to provide useful information than the one offered by intermediate peers. Therefore, in order to explore further the Vygotskian concept of zone of proximal development in physical activity contexts, future studies should attempt to include more ill-defined tasks which are known to promote more dialogue and multiple perspectives than tasks that are well defined (Cohen, 1994).

Although both genders displayed comparable feelings of self-efficacy, as
would support the gender-neutral connotation of the swimming task (Lirgg, 1991),
boys adopted higher ego goals than girls, who in turn scored higher on learning
goals, with respectively large and moderate effect sizes. These expected gender
differences are consistent with the results of earlier studies pertaining to goal ori-
entations in which it was found that females are more task-oriented than males,
who in turn are more ego-oriented (Duda, 1988; White & Duda, 1994). They also
support the notion of the most preferred source of competence information in late
adolescence (Horn, Glenn, & Wentzell, 1993), i.e., that females score higher on self-
comparison/ internal information sources and evaluative feedback from significant
others whereas males score higher on the use of competitive outcomes.

Although dispositional goal orientations were not assessed in the present
study, it can be speculated that the observed gender differences in goal involve-
ment indicate that participants focused on the characteristics of the context that were
most in agreement with their preferred orientation (Treasure & Standage, 1999).
It can be argued that boys were more likely to perceive the ego-involving charac-
teristics of the context (e.g., social visibility of the performance and presence of a
video camera) as salient whereas girls were more likely to perceive the context of
learning as a mastery climate (Solmon, 1996).

Consistent with our predictions, gender differences were also found on beha-
vioral variables. The swimming performance and number of task attempts were
higher for boys, while verbal information and demonstrations were more frequent
for girls, with all effect sizes being very strong (>70). These findings parallel earlier
research in the educational psychology literature conducted with children in the
areas of language (Ellis & Gauvain, 1992), reasoning (Tutge, 1992), or game playing
(Charlesworth & Dzur, 1987). This literature indicates that competitive interaction
and task implementation are predominant behaviors among boys whereas prosocial
and verbal behaviors prevail among girls.

The results of the present study suggest that social interaction modes identified
in childhood may be applicable to late adolescence. These results also extend previ-
ous findings in physical activity contexts which have demonstrated that although
boys and girls have similar activity levels in structured physical education, boys
are significantly more active than girls in unstructured recess periods (Sarkin,
McKenzie, & Sallis, 1997).

While some correlations between self-efficacy beliefs (i.e., level of self-effi-
cacy and self-efficacy for improvement) and performance support the predicting
function of self-efficacy for success (Bandura, 1977, 1997), the present study pro-
vides surprising findings about the relationship between self-efficacy and states of
goal involvement. Specifically, learning goals correlated negatively with self-effi-
cacy beliefs, the number of attempts, and performance, while ego goal involvement
correlated positively with self-efficacy and behavioral variables. These findings
differ from the results of previous experiments which have found that individuals
tend to have higher expectations when adopting self-referenced goals (Jourden,
Bandura, & Banfield, 1991; Schunk & Swartz, 1993).

It must be recognized that the cognitive variables and tools of previous stud-
ies were different from those of the present research. In earlier studies the goals
were considered as relatively stable tendencies, whereas in the present study they
were assessed as states related to a particular learning situation. Further, the results
indicated that the values for girls were in agreement with previous research whereas
the values for boys were not. One possible way to interpret the contrary relations
values for boys would be to see them in light of the protection of self-worth. When self-efficacy for a skill is high, ego goals can be adopted without being detrimental to achievement behaviors (Nicholls, 1984, 1989). This could be congruent with the fact that ego goals for boys were positively related to self-efficacy and behaviors. In contrast, when self-efficacy for a skill is low, social comparison is threatening to self-worth and can lead one to avoid challenging tasks (Nicholls, 1984, 1989).

In closing, we must acknowledge three important limitations of the present study. First, sampling size, the absence of a control group, and the use of single items to measure self-efficacy beliefs lead us to draw conclusions with caution. Second, the type of tutoring adopted in this study suffers from ecological validity. Specifically, the 8-minute one-on-one tutoring session of our experimental setting limits generalizability to the practitioner’s world. Future studies could therefore consider the role of tutor skill level and gender differences within procedures that are closer to typical physical education settings. For instance, longitudinal studies based on reciprocal peer tutoring, which gives student dyads the opportunity to alternate between peer student and peer teacher roles in the context of a structured format that guides them through the learning process (Fantuzzo et al., 1992), would warrant particular attention from sport psychology researchers and practitioners. The benefits of access to a large groups of same-age peers as compared to one-on-one tutoring (Schacter, 2000) could also be examined in the sport domain.

A third limitation of this study is that the measures retained to characterize the tutoring session were more focused on the tutor than on the tutee. Therefore, little is known about the tutee’s interaction in understanding the problem. Because this process is determined by the nature of the task and its definition (Cohen, 1994) and other factors such as age (Flavell, 1981) or skill level of the participants (Frayssée, 1985; Nelson-Le Gall, 1992), future research might conduct more qualitative inquiries and document other appropriate behavioral variables.

Despite these limitations, the increased understanding of factors that affect peer tutoring in motor learning contexts offered by this study has implications for coaching and physical education teaching practices. Evaluation of the competence gap between peers working in dyads seems important in understanding the emergence of particular psychological states, interactive dynamics between partners, and achievement. The present research suggests that novice learners in physical activity contexts, especially boys, could benefit from practicing in asymmetrical dyads using skilled peer tutors. Furthermore, the present study outlined gender differences in social learning conditions that may be informative to coaches and physical education teachers.

Further research should attempt to investigate the role of other personal characteristics of students such as achievement goal orientation or age, and the influence of the social relationship between peers (i.e., peer acceptance, friendship) on psychosocial variables, helping behaviors, and performance outcomes. Also, there is a need to explore the role of competence symmetry/asymmetry among more advanced learners. Furthermore, because Vygotsky’s zone of proximal development is a product of the interaction between both tutor and tutee, this limits how much a same-age tutor might help a tutee (Moll, 1990). Thus the benefits of cross-age peer-tutoring and the importance of training tutors could be studied in the context of physical activity.

Finally, although there has been little research in Vygotsky’s zone of proximal
development, much of the applied-behavior-analysis research on shaping behaviors of students (e.g., Cooper, Heron, & Heward, 1987; Miltenberger, 1997) is very similar work, and researchers may wish to determine the links if any. Together these research avenues should allow students in sport and physical education classes to participate more effectively in socially mediated learning experiences.

References


Thill, E., & Cury, F. (2000). Learning to play golf under different goal conditions: Their


Acknowledgment

The authors express sincere appreciation to Dr. Alan L. Smith and two anonymous reviewers for valuable feedback on previous drafts of this paper, and to Amélie Soulard for her assistance in the English translation. The authors also thank the principals, teachers, and students of the participating schools. A part of this paper was presented at the Annual North American Society for the Psychology of Sport and Physical Activity Conference, Clearwater Beach, FL, June 1999.

*Fabienne d’Arripe-Longueville is now with the UFRSTAPS, Université de Nice Sophia-Antipolis, 221, Route de Grenoble, 06200 Nice, France. Christophe Gernigon is now with the EA2991 Sport, Performance and Health, Université de Montpellier I, UFRSTAPS, 700 Ave. du Pic Saint Loup, 34090 Montpellier, France.*