Human Movement Science 66 (2019) 231–240

Contents lists available at ScienceDirect

Human Movement Science

journal homepage: www.elsevier.com/locate/humov

Manipulations to practice organization of golf putting skills through interleaved matched or mismatched practice with a partner

April Karlinsky, Nicola J. Hodges*

School of Kinesiology, University of British Columbia, 210-6081 University Boulevard, Vancouver, BC V6T 1Z1, Canada

ARTICLE INFO

Keywords:
Dyad practice
Contextual interference
Observational learning
Motor learning
Motor skills
Joint action

ABSTRACT

There is some evidence that alternating physical and observational practice with a partner for the same skill can benefit learning compared to practice alone. What has not been studied is whether a partner's interleaved practice impacts multi-skill learning, when the partner either matches or mismatches their partner's skill. Here we manipulated partners' practice schedules of two golf putting skills. Partners practiced the same ("matched") or different skills in alternation ("mismatched"). Based on previous research where interleaved demonstrations have induced beneficial contextual interference effects, we hypothesized that mismatching a partner on consecutive trials should also promote a similar type of interference in practice, which ultimately aids learning. A third control group was tested, where only one partner practiced while the other observed. All groups practiced for two days, with individual retention tests at the start of day 2 and one week later. Taking turns practicing and observing a partner did not benefit learning compared to the control, pure physical practice group and the matched and mismatched groups did not differ in outcomes. There was, however, evidence that partners were adapting their actions (i.e., compensating for over or undershooting of the target) based on the shots of their partner, in a similar manner to how they were adapting to their own errors. Thus, although partners were influencing each other's performance, it was not ultimately to the benefit (or cost) of overall learning. Partner-mismatching of skills through alternating practice was not sufficient to promote interference in practice and ultimately promote learning.

1. Introduction

People often practice motor skills in social settings, with similarly skilled peers. In these contexts, individuals not only engage in their own physical practice, but they also have the opportunity to observe others. Sometimes co-learners are practising the same skills as each other, or sometimes they are practising different skills. In this study, our aim was to determine if and how partners impact each other during practice when they are tasked with learning multiple skills. To address this aim, partners either repeated the same skill as their partner, across blocks of trials, or they performed a different skill to their partner. In this way, practice order was manipulated on a between-person level.

Taking turns with a partner can enhance the learning of a single skill compared to pure physical practice alone (e.g., Granados & Wulf, 2007; Shea, Wright, Wulf, & Whitacre, 2000; Shea, Wulf, & Whitacre, 1999; *cf.* Karlinsky & Hodges, 2018b). These interleaved

* Corresponding author.
E-mail address: nicola.hodges@ubc.ca (N.J. Hodges).

https://doi.org/10.1016/j.humov.2019.05.001
Received 20 August 2018; Received in revised form 30 April 2019; Accepted 2 May 2019
Available online 09 May 2019

periods of peer observation in between physical practice attempts are thought to promote information processing activities related to error-detection and strategy evaluation (e.g., Adams, 1986; Lee, Swinnen, & Serrien, 1994). When partners take turns practicing a single skill, the action they observe matches what they have just done and what they are about to do, such that the partner's action and associated outcome feedback are immediately relevant to the observer (e.g., Granados & Wulf, 2007; Karlinsky & Hodges, 2018b; Shea et al., 1999, 2000). It is unclear, however, how interleaved demonstrations impact learning in pairs when partners practice multiple skills. In these multi-skill cases, what the learner observes in a partner's trial may or may not be the same as their own next task.

Interleaved demonstrations that either match or mismatch the next action have been shown to impact the effectiveness of multi-skill learning in an individual learning context (Lee, Wishart, Cunningham, & Carnahan, 1997; Simon & Bjork, 2002). In one study, participants followed a blocked (repetitive) or randomly ordered practice schedule of three different keystroke timing tasks (Simon & Bjork, 2002). Consistent with the typical contextual interference (CI) effect, random groups performed worse in practice but more accurately in retention than blocked groups. However, if "correct" demonstrations (i.e., computerized displays) were given before each practice trial, these CI effects were substantially moderated. Demonstrations of the upcoming action that were different to the next to-be-practised skill (termed mismatched), compared to matched demonstrations, impaired acquisition but ultimately augmented retention for both the blocked and random groups. Here, demonstrations that promoted between-trial interference and cognitive effort enhanced learning. This is also congruent with evidence of CI effects following observational practice of high versus low interference practice schedules in others (e.g., Blandin, Proteau, & Alain, 1994; Wright, Li, & Coady, 1997).

In the keypress timing study detailed above (Simon & Bjork, 2002), benefits in retention were observed for mismatched demonstration groups despite lower perceptions of competence (judgments of learning) for the mismatched as opposed to the matched groups. There have been claims in recent studies that such competency perceptions are moderated by the practice context and can in some cases impact learning outcomes (e.g., Wulf & Lewthwaite, 2016). In dyad learning contexts, peer observation is thought to promote higher competency perceptions in co-learners, in comparison to practice alone, as watching similar others perform successfully can help observers believe they are also capable of achieving similar levels of performance (Bandura, 1977; Lewthwaite & Wulf, 2012). Observing a peer is also thought to promote individuals' motivation to improve, potentially by adding a sense of competition to the practice context (e.g., Granados & Wulf, 2007; McNevin, Wulf, & Carlson, 2000). Thus, measures of motivation and competency are needed in studies of paired practice to determine if practice with a partner impacts these variables, whether this depends on the schedules each partner adopts and ultimately, whether this might explain any learning-related benefits that ensue.

There is evidence that individuals practising alone are biased by their own previous motor responses. For example, when aiming in alternation to targets with amplitudes of 20° and 60°, people tended to overshoot the shorter distance target and undershoot the longer distance target (e.g., Sherwood, 2010; Sherwood & Fosler, 2013; Sherwood & Rothman, 2011). In paired action contexts, there is also evidence that individuals monitor a partner's performance and adapt their own behaviours in response to observed errors in a peer, similar to as if they were their own (e.g., de Bruijn et al., 2012; de Bruijn, Miedl, & Bekkering, 2011; *cf.* Picton, Saunders, & Jentzsch, 2012). For example, individuals have been shown to exhibit "post-error slowing" following both their own and a co-actor's errors in reaction time tasks (e.g., de Bruijn et al., 2011, 2012). In a sequence learning task, where individuals were given choice over when to switch between practice of three different sequences, the interleaved practice behaviours of a partner influenced their frequency of switching between different skills (Karlinsky & Hodges, 2018a). In a subsequent study, partners' decisions to repeat or switch between different skills were also shown to be influenced both by their own errors as well as their partner's errors (Karlinsky, Alexander, & Hodges, in preparation). Therefore, examination of practice behaviours on a trial-to-trial basis, both within and between individuals, offers insights into how partners can impact the practice of co-learners.

Here we report a novel dyad practice protocol wherein we manipulated the relationship between what is observed and physically practiced on consecutive trials. A golf putting task was used with two different putters requiring standing or seated putting to the same target distance. Partners either used the same putter as their partner (matched group) or used a different putter on consecutive trials (mismatched group). In the former case, there was a match between what was observed and performed, minimizing between-person interference and making practice easier, but potentially at the cost of learning. However, in this matched case, there would be immediate opportunities to learn from a partner's performance to aid with execution of the upcoming skill and enhance perceptions of competency, which might also aid learning.

Trial-to-trial errors were measured in order to determine whether there was evidence of "post-error compensation" between partners. Here, over- or undershooting behaviours in the partner would be "corrected" for by the other partner as determined by an opposite pattern of over- or undershooting. We expected any compensation-type behaviours between partners to be higher in the matched versus mismatched pairs, because of the similarities between the skills and the potential attention given to skills immediately relevant to current performance.

We additionally tested a control dyad group, where partners either just practiced or just watched, thus controlling for the presence of a partner (for a review of "audience effects", see Hamilton & Lind, 2016). We anticipated that physical practice would enhance learning compared to observational practice (e.g., Blandin et al., 1994; Karlinsky & Hodges, 2014; Shea et al., 2000; Wright et al., 1997), but that alternating practice with a partner would be more effective than pure physical practice (e.g., Granados & Wulf, 2007; Shea et al., 1999, 2000). The IMI (Intrinsic Motivation Inventory, n.d.) was used to determine whether practising with a partner was more motivating than practising alone (Lewthwaite & Wulf, 2012) as well as to provide information about competency perceptions. The IMI provides subscale information about interest/enjoyment, competence, choice/autonomy, and pressure/tension. We hypothesized that the alternating pairs would rate their interest/enjoyment of the practice experience more positively than individuals engaging in just physical practice. It is possible, however, that taking turns with a partner would also augment the pressure/tension experienced than the more independent physical practice control condition (e.g., Rhea, Landers, Alvar, & Arent, 2003) and that mismatched practice would be associated with lower perceptions of competence than matched (see Simon & Bjork, 2002). Participants additionally responded to a customized paired practice experience questionnaire, probing their desire to outperform their partner (see McNevin et al., 2000) as well as other perceptions of the practice experience.

2. Methods

2.1. Participants

Seventy-six right-handed females (M = 22.1 yr, SD = 5.3) volunteered to participate and were paid $10/hr. We chose to restrict participant eligibility to females in order to limit any potential within-dyad (e.g., Carli, 1989; Harskamp, Ding, & Suhre, 2008) or participant-experimenter sex-related effects (e.g., Rumenik, Capasso, & Hendrick, 1977), as a female experimenter conducted the data collection. There is also some evidence that joint (dyadic) action effects are larger for same-sex pairs compared with opposite-sex pairs (Mussi, Marino, & Riggio, 2015; van der Weiden, Aarts, Prikken, & van Haren, 2016). All participants had normal or corrected-to-normal vision, no known neurological or motor disorders, and limited golf putting experience. Informed consent was obtained, and the study was conducted in accordance with the Behavioural Research Ethics Board of the University.

2.2. Groups and partners

Participants were strangers and paired based on chance and availability. Pairs were assigned to either the matched, mismatched, or control group (n = 24 per group; 12 pairs/group). Two additional pairs were tested but were excluded post-testing due to experimenter mistakes in the order of testing and participant ineligibility due to previous golf experience. Within the matched and mismatched groups, one member of each pair was randomly assigned to be Partner 1 (P1), which indicated that they would perform the first trial (and all subsequent odd number trials) of the acquisition sessions. The other member of the pair was assigned to be Partner 2 (P2), which meant that they would perform the second trial (and all subsequent even number trials) of the acquisition sessions. We manipulated the putter schedule of the matched and mismatched group P1s such that the selected putter (standard or miniature putter) would serve as a demonstration of the upcoming putt for P2s in the "matched" group, but be a demonstration of a different putt, with a different putter, for the "mismatched" group P2s. All P2s practised in the same practice order regardless of group. Importantly, both partners within a group practised under the same matched or mismatched conditions in a semi-blocked order of trials, except that P2s matched (or mismatched) their partner's task (i.e., putter) on every trial. Because P1s went first, on the first trial of every block of new trials they did not match (or mismatch) the P2's previous task (14% of trials).

Within the control group, one member of each pair was assigned to be the "Actor", who physically practiced during the acquisition sessions, while the other partner was assigned to be the "Observer", who watched the Actor's acquisition sessions. Actors completed the same order of trials and putters as Partner 2 s.

2.3. Task and apparatus

Participants had to putt a golf ball so that it stopped in the centre of a target. Two different right-handed putters were used: a "standard" putter (91 cm) and a "miniature" putter (40 cm). The standard putter was used standing, sideways to the target, and the miniature putter was used seated, facing the target. Standard white golf balls were used throughout the study. Putts were made on a flat green felt surface (2 m wide × 7 m long) to a target (black cross) 150 cm away. Participants wore Plato liquid crystal occlusion glasses (Translucent Technologies, Ontario) and ear protectors (3 M Optime 101) during the no-feedback (FB) pretest and retention tests.

2.4. Materials

We used two questionnaires. The Task Evaluation Questionnaire (TEQ) is a 22-item version of the Intrinsic Motivation Inventory designed to assess interest and attitudes toward an experimental task (Intrinsic Motivation Inventory, n.d.).[1] The TEQ consists of four subscales: interest/enjoyment (7 items), perceived competence (5 items), perceived choice (5 items), and pressure/tension (5 items). The wording was modified for golf putting and items were answered on a 7-point Likert scale (1 = *not at all true* to 7 = *very true*). We also had a customized paired practice experience questionnaire which included three questions (detailed in the Results, Table 3) which were rated on the same 7-point Likert scale for truthfulness.

2.5. Procedural design

A summary of the study design is provided in Table 1. The study was conducted over three days, with days 1 and 2 comprising acquisition sessions which were completed in pairs and "day 3" (one week later),completed alone. Day 1 also consisted of a 6-trial familiarization phase and 6-trial, no feedback pretest, which were completed alone (3 attempts/putter, with order counterbalanced

[1] The Task Evaluation Questionnaire and additional variations of the Intrinsic Motivation Inventory can be accessed for scholarly use by registering with the Self-Determination Theory Organization through the link provided in the reference list (Intrinsic Motivation Inventory, n.d.).

Table 1
Experimental design and procedures.

	Day 1		Day 2 (~24 h later)			Day 3 (~1 week later)	
	Pretest	Acquisition	Retention	Retention	Acquisition	Retention	Retention
Alone or Paired	Alone	Paired	Alone	Alone	Paired	Alone	Alone
Feedback	X	✓	X	✓	✓	X	✓
# Trials/Putter	3	18	6	6	18	6	6
# Trials/Partner	6	36	12	12	36	12	12

Note. During the paired acquisition sessions, partners alternated turns on consecutive trials using the same putter (matched group) or different putters (mismatched group), or the Actor physically practiced while their Observer partner watched (control group).

for each). To prevent any learning in the pretest, participants wore occlusion glasses and ear protectors, such that outcome feedback was withheld.

In the paired acquisition phase (completed on day 1 and day 2) for the matched and mismatched groups, partners alternated turns for a total of 72 trials (36 physical practice trials and 36 observation trials per partner). Partner 1 (P1) putted on trial 1 and Partner 2 (P2) using the same putter (matched group) or different putter (mismatched group), putted on trial 2 etc. Whether P1 used the miniature or standard putter first was counterbalanced. Although partners alternated turns on each trial, each participant followed a semi-blocked physical practice schedule; using their assigned putter for six trials before switching to the other putter. This resulted in six physical practice blocks per partner (3 blocks/putter). In the control group, the Actors were exactly matched to the putter practice order of the P2s.

Day 2 of the study took place ~24 h later and began with two retention tests conducted alone; without feedback (identical to the pretest) and then with feedback. Each retention test was 12 trials, comprising blocks of three trials with the same putter before switching to the other putter (order counterbalanced). The second acquisition phase was conducted immediately after retention testing. Day 3 of the study took place ~1 week after day 2 and was completed individually. Participants completed the same two retention tests as performed on day 2.

2.6. Procedure

On day 1 of the study, partners introduced themselves and were allowed time to chat. They were informed that their goal was to learn how to hit the golf ball with two different putters, so that it landed as close as possible to the target centre. Instructions on how to hold the standard and miniature putters were provided and the experimenter demonstrated the general technique without hitting a ball.

Participants began by completing a familiarization phase, followed by the no-feedback pretest. The experimenter manually occluded the occlusion glasses upon ball-strike. The ball's distance from the target was recorded and the ball was removed before restoring the participant's vision. Partner 1 (P1) always performed the familiarization and pretest first, while P2 responded to questions regarding their previous golf-related experiences (e.g., number and approximate dates of miniature golf, 9-hole, and 18-hole golf experiences)[2] and a handedness questionnaire (Oldfield, 1971) in another room. These questionnaires were used to confirm the eligibility criteria of limited golf experience and right-hand dominance.

In acquisition, partners in the matched and mismatched groups were instructed to watch one another's practice from a designated observation spot (from the side and near the target so they could see both the putting action from the side and the ball's outcome), but not to communicate about the putting tasks. They were encouraged to pay attention to their partner's practice trials to help them learn, given that they would be tested using both putters the following day/week. Outcome feedback was given on every trial.

In the control group, Observers watched their Actor partner practice from the same observation location as the pairs in the matched and mismatched groups. Observers were also encouraged to attend to the Actor's practice and told they would also be tested on the putting skills the following day/week. The Actors' inter-trial breaks were controlled to match the timing of the other two groups. As such, Actor controls only differed from the other two groups in terms of the opportunity to observe their partner between practice attempts and the putter order was exactly matched to that of the Partner 2 s.

Day 2 of the study began with two retention tests conducted alone. The no-feedback retention test was always completed first to allow us to assess how well the golf putting skills were retained while preventing further learning. The subsequent, with-feedback test matched the conditions of practice and provided an index of participants' capabilities without the potentially perturbing effects associated with the removal of vision. Participants then completed another period of acquisition, identical to day 1. Afterwards, all participants (except the Observers) completed the Task Evaluation Questionnaire (Intrinsic Motivation Inventory, n.d.), responding with respect to how they generally felt during both paired practice sessions.

On day 3 of the study, participants completed the same two retention tests as performed on day 2 and responded to a paired

[2] As Partner 2 s always completed the golf experience questionnaire before completing the pretest, it is possible that this opportunity to reflect upon their previous golf-related experiences could have affected their perceptions of self-efficacy, and as a potential corollary effect, their performance and learning. However, given the lack of partner-related effects throughout the experiment, this was unlikely the case in the current study.

practice experience questionnaire, which consisted of three questions about the practice experience (full text questions are presented in Table 3). Participants were then fully debriefed and compensated for their time.

2.7. Measures and analysis

2.7.1. Error

The distance of the ball to the target centre was measured as constant error (CE) in the x- and y-dimensions after every trial. These data were transformed into radial error (RE) and bivariate variable error (BVE) for two-dimensional measures of accuracy and consistency, respectively (see Hancock, Butler, & Fischman, 1995). RE was calculated for each trial based on the CE in the x- and y-dimensions using Pythagoras' theorem: $RE = (x^2 + y^2)^{1/2}$. BVE was calculated as a standard deviation (SD) measure across two dimensions, given by the square root of a participant's k shots' mean squared distance from their centroid (x_c, y_c), where the centroid is defined as the positionally typical shot (average x and y position) for a participant's block of trials: $BVE = \{(1/k)[\sum_{i=1}^{k} (x_i - x_c)^2 + (y_i - y_c)^2]^{1/2}\}$.

Our primary analyses were on testing for differences between the matched and the mismatched groups (n = 24/group). Secondary analyses were conducted to compare the Actors who did not alternate with a partner to the Partner 2 s from the matched and mismatched groups (n = 12/group). We also compared the Actors to the Observers, mostly for descriptive purposes, to determine any benefits from pure observation in this golf putting task (the absence of a no-practice control group prevented a strong test of this question).

Errors for the matched and mismatched groups were compared in 2 Group (matched, mismatched) × 2 Partner (P1, P2) × 2 Putter (standard, miniature) mixed ANOVAs, with repeated measures (RM) on the last factor. We mostly included Putter as a blocking variable as participants were more accurate with the standard than miniature putter in all phases of testing. Only relevant group interactions for putter are discussed. Separate analyses were conducted for each error measure (RE and BVE), for the Pretest, Acquisition, and Retention phases.

Acquisition data were grouped into 6-trial blocks (one putter/block), resulting in an additional RM factor of Block. Linear trend analysis was conducted on this Block factor. In retention, an additional RM factor of outcome feedback was included. In secondary analyses, comparisons were also made across the P2s and Actors only, and across the Observers and Actors. The same tests as above were used, with the group factor reflecting these additional comparisons.

2.7.2. Behaviours

Compensatory (or corrective) behaviours were defined and measured in practice. These behaviours were first determined based on whether an overshoot or undershoot of the target in the y-dimension resulted in compensation (i.e., a shorter or longer putt, respectively) on the subsequent trial for that individual (self-referenced correction) and/or for their partner (partner-referenced correction). These data (percentages) were analyzed in a 2 Group (matched, mismatched) × 2 Partner (P1, P2) × 2 Day (Day 1, Day 2) × 2 Correction-type (self-referenced, partner-referenced) mixed ANOVA, with RM on the last two factors. For comparison, we also analyzed the percentage of self-referenced corrections of the Actor group and compared these statistically to the Partner 2 s from each of the dyad groups, in a 3 Group × 2 Day mixed ANOVA.

2.7.3. Questionnaires

We used Cronbach's alpha to assess the internal consistency of the items within each subscale of the Task Evaluation Questionnaire (TEQ), to provide an index of each subscale's reliability. For the current sample, Cronbach's alpha values were good for the interest/enjoyment ($\alpha = 0.93$), perceived competence ($\alpha = 0.86$), and pressure/tension ($\alpha = 0.86$) subscales; however, reliability for the perceived choice subscale was weak ($\alpha = 0.69$), so we do not report these data. Participants' average scores for each subscale of the TEQ were submitted to separate ANOVAs. We first tested for differences between the matched and mismatched alternating groups using separate two-way ANOVAs for each TEQ subscale, with Group (matched, mismatched) and Partner (P1, P2) as the between-factors. Similar analyses were run on the paired practice experience questions. Secondary analyses of the TEQ subscales were performed comparing the Actors to the matched and mismatched Partner 2 s for each subscale (one-way ANOVAs).

For all analyses Greenhouse-Geisser corrections were applied to the degrees of freedom for violations to sphericity. Significant effects and interactions were followed up with Tukey's HSD procedures (all $ps < 0.05$ reported). Partial eta squared (η_p^2) and Cohen's d values are reported as measures of effect size for ANOVAs and t-tests, respectively. Power values ($1 - ß$) are given for non-statistically significant effects where $F > 1$. We only report full results when $F > 1.5$.

3. Results

3.1. Error

To summarize the results for error, although there was evidence of improvement in groups, there were no significant group differences in acquisition or retention. Detailed analyses for relevant comparisons are given below.

3.1.1. Pretest

The matched and mismatched groups' mean RE (indexing accuracy) and BVE (indexing consistency) in testing are presented in

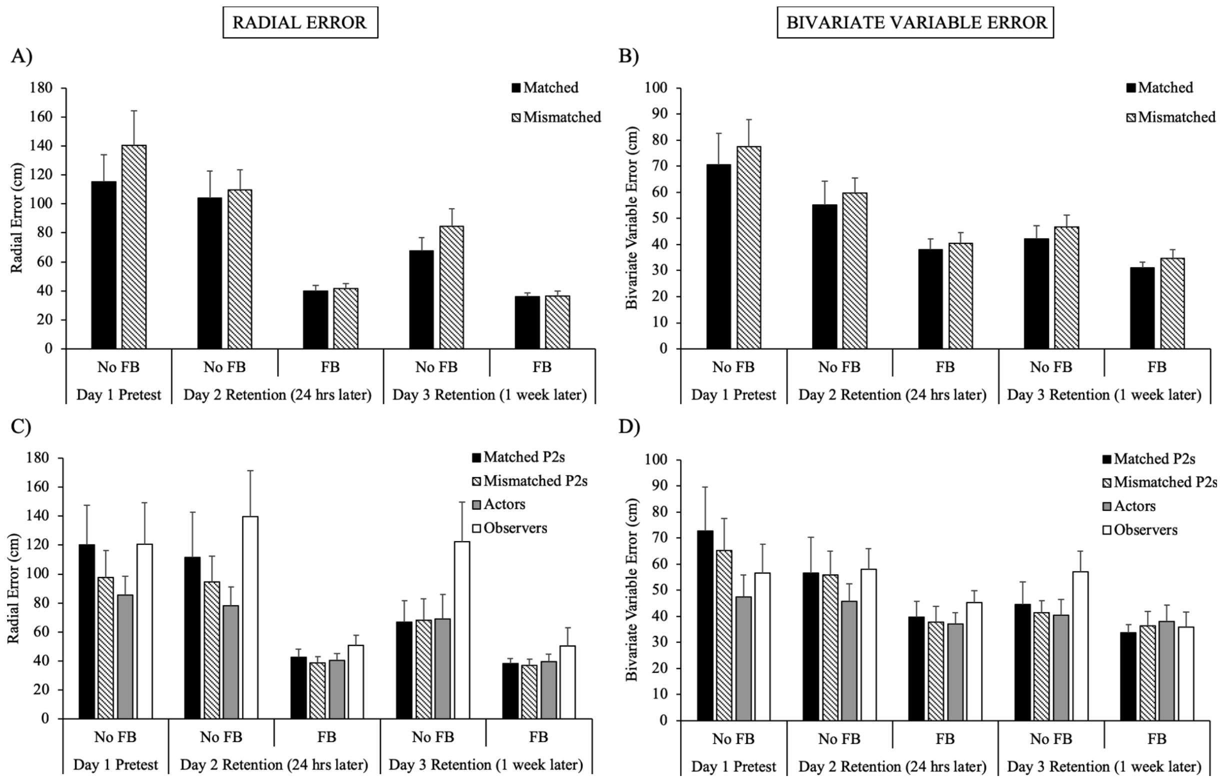

Fig. 1. A-D. Pretest and retention testing mean radial error (1A) and mean bivariate variable error (1B) for the Matched and Mismatched alternating dyad groups (n = 24/group) and the Matched and Mismatched Partner 2 s (P2s) and control group Actors and Observers (1C & 1D, n = 12/group). FB = feedback (error bars = SE).

Fig. 1A and 1B (n = 24/group). The groups did not differ on RE ($F = 1.1$; Fig. 1A). There were also no Partner differences, $F(1, 44) = 2.41$, $p = .13$, $1 - ß = 0.33$; but the Group × Partner interaction was close to accepted significance levels, $F(1, 44) = 3.86$, $p = .056$, $1 - ß = 0.48$. In the mismatched group, there was a trend for P2s to be more accurate than P1s (data for Partner 2 s only are shown in Fig. 1C and 1D). The BVE data (Fig. 1B) showed the same pattern of results (no differences between groups or partners with respect to consistency, $Fs < 1$; Group × Partner, $F(1, 44) = 1.67$, $p = .20$, $1 - ß = 0.24$).

Comparisons between the Actor group and Partner 2 s from the matched and mismatched groups (n = 12/group) showed no initial between-group differences in RE ($F = 1.4$) or BVE, $F(2, 33) = 1.73$, $p = .19$, $1 - ß = 0.34$ (see Fig. 1C and 1D). Similarly, as illustrated in Fig. 1C and 1D, comparisons of the Actors with the Observers showed no differences in pretesting in terms of RE, $F(1, 22) = 2.06$, $p = .17$, $1 - ß = 0.28$, or BVE ($F < 1$).

3.1.2. Acquisition and retention

As illustrated in Fig. 2A and 2B, participants in the matched and mismatched groups improved their accuracy (RE) and

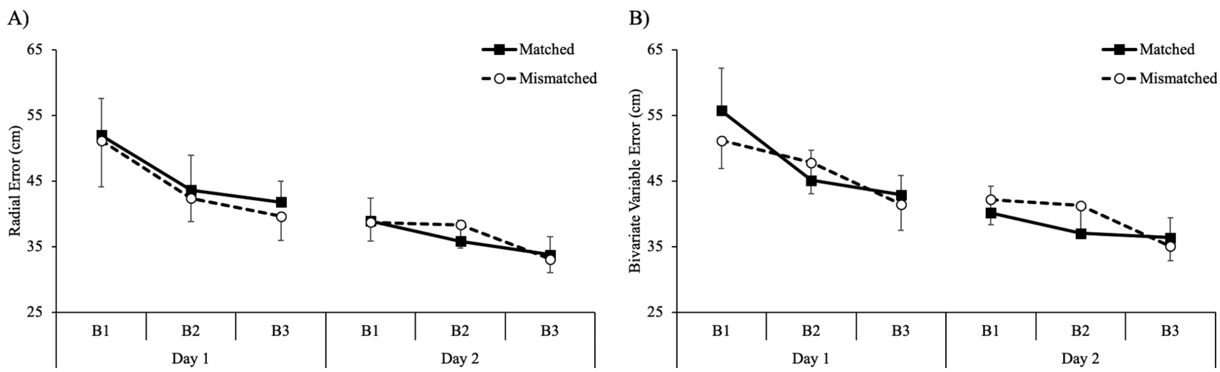

Fig. 2. A & B. Acquisition mean radial error (2A) and bivariate variable error (2B) for the Matched and Mismatched alternating dyad groups. All blocks consisted of 6 trials/putter (error bars = SE).

consistency (BVE) across blocks on day 1, confirmed by linear trends ($ps < 0.001$), and for RE but not BVE on day 2 ($p = .004$). However, the groups did not differ on either day of practice ($Fs < 1$). There were also no partner-related effects during practice, with respect to putting accuracy (RE: $Fs < 1$) or variability (BVE: day 1, $F(1, 44) = 2.69$, $p = .18$, $1 - ß = 0.36$), nor any Group × Partner interactions ($Fs < 1$).

There were also no group-related differences in retention, as illustrated in Fig. 1A and 1B. For RE on day 2, all $Fs < 1$ (Group × Partner, $F(1, 44) = 1.52$, $p = .22$, $1 - ß = 0.23$). For BVE, all $Fs < 1$. Having outcome feedback aided accuracy (lower RE), $F(1, 44) = 44.44$, $p < .001$, $\eta_p^2 = 0.50$, and consistency (lower BVE), $F(1, 44) = 13.92$, $p = .001$, $\eta_p^2 = 0.24$. When tested one week later on "day 3", there were again no group- or partner-related retention effects with respect to RE or BVE (Fs ranged from < 1 to 1.4).

Secondary analysis of the Actor group, in comparison to the Partner 2 s from the matched and mismatched dyad groups, also failed to show differences in acquisition (not shown) and retention (Fig. 1C and 1D). Only on day 1 of practice were the group-related F values > 1.5, due to trends for the Actor control group to perform with greater accuracy (lower RE), $F(2, 33) = 2.91$, $p = .069$, $1 - ß = 0.53$, and consistency (lower BVE), $F(2, 33) = 2.77$, $p = .077$, $1 - ß = 0.51$.

Comparisons of the Actors to the Observers showed only trends for the Actors to be more accurate and consistent. As shown in Fig. 1C and 1D, on day 2 retention, the Actors had lower RE and BVE, although neither difference was statistically significant, RE: $F(1, 22) = 3.93$, $p = .060$, $1 - ß = 0.48$, BVE: $F(1, 22) = 3.43$, $p = .077$, $1 - ß = 0.43$. Similar trends were seen the following week, although again there were no significant group differences for both RE, $F(1, 22) = 2.74$, $p = .11$, $1 - ß = 0.35$, and BVE, $F(1, 22) = 1.38$, $p = .25$, $1 - ß = 0.20$. The only other group difference was a Group (Role) × Putter interaction for BVE, $F(1, 22) = 5.07$, $p = .035$, $\eta_p^2 = 0.19$. Actors were more consistent than Observers when using the miniature putter.

3.2. Behaviours

As illustrated in Fig. 3A, all participants showed a high frequency of compensatory behaviours (~70–76% of trials). The frequency of corrections was similar for self-referenced corrections (increasing or decreasing the shot if the target was under- or over-shot on the person's previous trial) and partner-referenced corrections (increasing or decreasing the shot if the target was under- or over-shot on the partner's preceding trial). Statistical analysis did not yield differences between self- and partner-referenced compensations ($F < 1$). Contrary to expectations, there were also no group-related effects (all $Fs < 1$ with the exception of Group × Day × Correction-type, $F(1, 44) = 2.07$, $p = .16$, $1 - ß = 0.29$). There was an effect of Partner for this analysis, with the Partner 2 s, who followed Partner 1 s, compensating more frequently ($M = 74\%$, $SD = 7\%$), than their leading Partners ($M = 71\%$, $SD = 14\%$), $F(1, 44) = 6.34$, $p = .016$, $\eta_p^2 = 0.13$. Although there was an interaction with Correction-type (self or partner), $F(1, 44) = 6.43$, $p = .015$, $\eta_p^2 = 0.13$, this was not a result of more partner-related compensations. Rather, P2s made more self-referenced corrections ($M = 76\%$ of trials, $SD = 7\%$) than P1s ($M = 68\%$, $SD = 11\%$).

When looking at the percentage of compensation behaviours for the Actors, a similar percentage of corrections was seen as noted for the dyad groups ($M = 74\%$ of all acquisition trials, $SD = 7\%$). Statistical comparisons of the Actors with the Partner 2 s did not yield group differences ($Fs < 1$; Group × Day: $F(2, 33) = 1.21$, $p = .31$, $1 - ß = 0.25$; data not shown).

Because such corrective behaviours cannot definitively be attributed to a person's preceding physical practice trial or an observed partner's trial, we conducted a descriptive analysis to help further determine whether corrections were likely self- or partner-referenced. Four types of corrections were identified as illustrated in Fig. 3B. Putts were classified based on the presence of a corrective

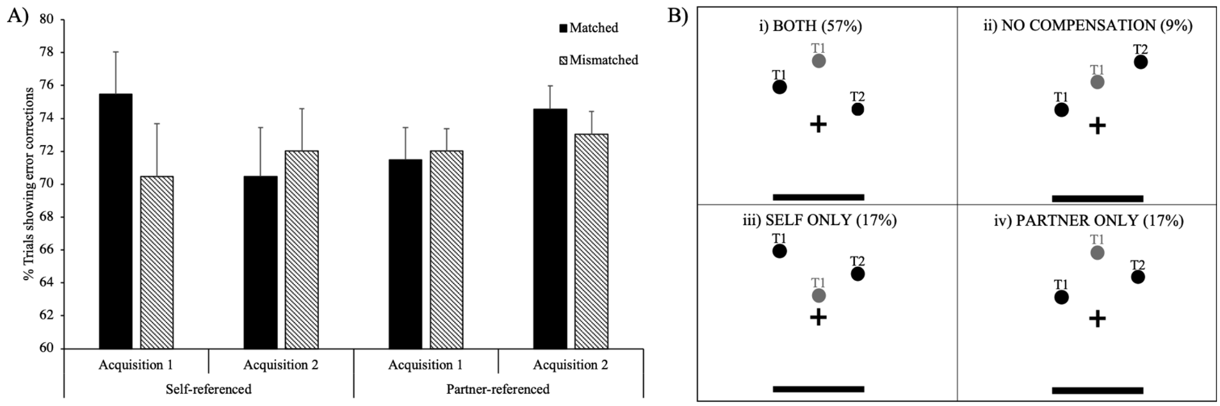

Fig. 3. A & B. Mean percentage of acquisition trials where Matched and Mismatched alternating dyads demonstrated corrective behaviour in response to y-dimension error in the preceding performed (self-referenced) or observed (partner-referenced) trial (3A). Schematic of potential corrective behaviours to errors in the y-dimension (i.e., overshooting or undershooting the target), along with percentage of acquisition trials on which the Matched and Mismatched dyad groups showed each type of scenario (collapsed across groups and days, 3B). Examples pertain to the outcome of 'trial 2' ("T2") with reference to the P2's own previous trial (black "T1") or their partner's previous trial (grey "T1"). Outcomes could reflect i) compensation with respect to both their own and their partner's previous error, ii) no compensatory behaviour, iii) compensation with respect to only their own previous error, or iv) compensation with respect to only their partner's previous error.

Table 2
Mean ratings (and SDs) for the Task Evaluation Questionnaire subscales.

Subscale	Matched		Mismatched		Actors
	P1s	P2s	P1s	P2s	
Interest/enjoyment	5.2 (1.5)	4.8 (1.1)	4.3 (1.4)	4.6 (1.1)	4.7 (1.3)
Competence	3.9 (1.1)	3.3 (1.2)	3.7 (1.1)	3.8 (1.0)	3.7 (1.0)
Pressure/tension	2.1 (0.9)	2.5 (1.0)	2.3 (1.3)	2.8 (0.8)	2.1 (1.1)

Note. Scales ranged from 1 to 7.

behaviour with reference to both the person's own and their partner's previous trial (i), the absence of a corrective behaviour with reference to either the person's own or their partner's previous trial (ii), the presence of corrective behaviour with reference to only the person's own previous trial (iii, self-only), or corrective behaviour with reference to only the partner's previous trial (iv, partner-only). The vast majority of trials showed compensatory behaviours which could be classed as both self- and partner-referenced (57%). The percentage of trials which showed only partner-referenced corrections or only self-referenced corrections, although markedly smaller, did not differ (17% for each).

3.3. Questionnaires

Mean responses to the Task Evaluation Questionnaire subscales are presented in Table 2. Practicing under matched versus mismatched conditions did not differentially impact perceptions. Both groups generally reported medium to high ratings for interest/enjoyment and competence, and low ratings for perceived pressure/tension (interest/enjoyment: Group, $F(1, 44) = 1.91$, $p = .17$, $1 - ß = 0.27$; pressure/tension: Partner, $F(1, 44) = 2.53$, $p = .12$, $1 - ß = 0.34$; all other Fs < 1.1). There were also no significant differences between the Actors and the Partner 2 s from the alternating dyad groups (Fs < 1).

The paired practice experience questionnaire items and results are presented in Table 3. Regardless of whether schedules were matched or mismatched, watching a partner's practice was rated as moderately helpful, Group, $F(1, 44) = 1.83$, $p = .18$, $1 - ß = 0.26$, and not interfering, Group, $F(1, 44) = 2.11$, $p = .15$, $1 - ß = 0.30$, for participants' own performance/learning. In terms of the desire to be more accurate than their partner, ratings were relatively high and again these did not vary as a function of group (all Fs < 1).

4. Discussion

Our primary aim was to assess if and how practice with a partner who engages in the same (matched) or different order of trials (mismatched) impacts motor learning. We hypothesized that alternating turns with a partner who practices a different skill (mismatched) should promote between-person, observation-induced interference, which would facilitate retention in comparison to matched-skill practice (Lee et al., 1997; Simon & Bjork, 2002). However, contrary to our hypotheses, the matched and mismatched alternating practice groups showed essentially the same outcomes, in terms of errors in acquisition and retention and perceptions of the practice experience. Although we discuss potential reasons for the lack of difference below, it may just be that observation-induced interference in the form of alternating practice for novices is not sufficient to bring about differences in performance and learning when the actual physical practice experiences for these participants remains the same across groups.

In addition to a lack of difference between groups that alternated turns practicing and observing, these groups did not outperform an Actor group who essentially practiced alone with a passive observing partner. This conflicts with results from some previous dyad learning literature, where alternating practice enhanced the learning of a single skill compared to practice alone (e.g., Granados & Wulf, 2007; Shea et al., 1999, 2000; yet see Karlinsky & Hodges, 2018b). Given that we controlled for inter-trial intervals in this design as well as potential social facilitation/audience effects associated with having a partner watch, suggests that previous benefits of dyad practice may be related to these inter-trial rest or social facilitation factors.

Although there were no outcome-related differences between groups, there was evidence that pairs were sharing in the practice experience. Looking at the alternating behaviours of the dyads, there was evidence of compensation between and within individuals. Overshooting (or undershooting) the target was equally likely to lead to a shorter (or longer) putt on the next turn regardless of

Table 3
Mean ratings (and SDs) for the paired practice experience questionnaire.

Item	Matched		Mismatched	
	P1s	P2s	P1s	P2s
1. Watching my partner helped my own performance/learning	3.9 (1.5)	4.7 (1.2)	3.6 (1.8)	3.8 (1.5)
2. Watching my partner interfered with my own performance/learning	1.8 (1.0)	2.8 (1.7)	2.9 (1.8)	3.1 (1.7)
3. I wanted to be more accurate than my partner	5.0 (1.3)	5.2 (1.4)	5.0 (1.7)	4.8 (1.5)

Note. Scales ranged from 1 to 7.

whether it was in response to the person's own previous trial or their partner's. These results are in line with research showing that co-actors monitor each other's performance and adapt their behaviours in response to a partner's errors and/or task constraints, similarly to as if they were their own. For example, in addition to post-error slowing (e.g., de Bruijn et al., 2011, 2012), individuals have been shown to adjust their own movement kinematics (to lift their arm higher) when they know a co-actor needs to clear an obstacle (Schmitz, Vesper, Sebanz, & Knoblich, 2017; van der Wel & Fu, 2015).

The addition of the Actor group that did not alternate, but was observed, allowed us to compare physical versus observational practice for these golf putting skills. As anticipated, the Actors generally outperformed the Observers on the retention tests, consistent with previous research (e.g., Blandin et al., 1994; Karlinsky & Hodges, 2014; Shea et al., 2000; Wright et al., 1997). The results also suggest that observational practice aided skill development, in view of the similarity between the Actors' and Observers' performance on the final, with-feedback retention test. This is in line with research showing that benefits of observation are not fully appreciated until individuals have had the opportunity to receive knowledge of results (KR) on their own performance (which the Observers received during the day 2 with-feedback retention test; e.g., Andrieux & Proteau, 2013; Deakin & Proteau, 2000). Although the inclusion of retention tests both with and without KR allowed us to assess learning for observers once they received feedback on their own performance, this did render the ratio of practice to retention trials relatively high. Even though learning should have been prevented on the no-KR trials and the KR trials allowed us to assess learning specific to the conditions of practice, larger final differences between groups (especially Observers and Actors) might have been expected if these test trials had been limited.

It is important to consider potential factors that differed between the current study and previous research, which may have contributed to the pattern of results and specifically the lack of group differences in outcomes. Here we used relatively complex golf putting skills, as opposed to relatively simple keystroke timing tasks which have yielded the strongest contextual interference (CI) effects and influence of observation or interleaved demonstrations on learning (Lee et al., 1997; Simon & Bjork, 2002). It may be that interference-enhancing protocols (such as mismatched demonstrations) impact more on simpler, laboratory-type tasks, as compared to more complex, applied skills, because the former are intrinsically less interesting and challenging (e.g., Guadagnoli & Lee, 2004; Lee & White, 1990; Wulf & Shea, 2002). Observation-induced interference effects, especially as brought about by the schedule of another person, might only impact on individual motor learning for relatively simple and unchallenging tasks.

Another difference and potential explanation for the equivocal effects of matched versus mismatched practice is that here we used a real person as the model (i.e., a co-learner), who was also learning the task and making errors, as compared to the computer-based models of perfect performance used to-date (Lee et al., 1997; Simon & Bjork, 2002). It is possible that learning models in this multi-skill learning protocol brought too much interference into practice, hindering learning of this task. However, the lack of difference in practice between groups based on matched or mismatched schedules does serve to question this explanation. Moreover, the mismatched group did not show higher ratings in terms of their perceptions of the degree of interference experienced during practice as compared to the matched group. Neither did the mismatched group perceive watching their partner to be less helpful than reported by the matched group. It may be useful in the future to pair individuals with more (or less) experienced peers to better manipulate this potential variable of amount of between-person interference.

In summary, we studied multi-skill practice organization in pairs, where partners practiced the same or different skills in alternation (n = 24/group for the dyad pairs and n = 12/group for Actors and Observers). Mixed physical and observational practice with a partner did not enhance (or impede) learning compared to pure physical practice, nor did matched versus mismatched practice schedules with a partner differentially impact skill acquisition. Given that partners bring more variability into practice than "perfect" models of an upcoming skill, it might be that the degree of interference was too high to benefit learning in this task. However, we had no evidence of interference in practice for the mismatched group in terms of outcome error (as compared to the matched dyads), nor higher perceptions of interference. Because there was no group that had higher or lower within-person CI, we are not able to say whether this task (i.e., golf putting with 2 skills) would be sensitive to traditional CI effects and as such make comparisons between within- and between-person CI. This will be important for future studies as we evaluate these dyad practice conditions. Considering the inter-trial breaks commonly incorporated into applied training settings for a variety of reasons, it will be important to continue such inquiries into how to organize paired multi-skill practice to optimize learning. However, on the basis of these data, we have no evidence that partner-interleaved practice serves to bring between-person CI into practice, or if it does, such observation-induced interference is not sufficient to enhance learning in such semi-blocked practice schedules as adopted here.

Declaration of Competing Interest

None.

Acknowledgements

Funding: This work was supported by the Natural Sciences and Engineering Research Council of Canada [2016-0426].
We would like to thank Jaspreet Dhillon for her assistance with data collection.

References

Adams, J. A. (1986). Use of the model's knowledge of results to increase the observer's performance. *Journal of Human Movement Studies, 12*(2), 89–98.
Andrieux, M., & Proteau, L. (2013). Observation learning of a motor task: Who and when? *Experimental Brain Research, 229*(1), 125–137.
Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215.

Blandin, Y., Proteau, L., & Alain, C. (1994). On the cognitive processes underlying contextual interference and observational learning. *Journal of Motor Behavior, 26*(1), 18–26.
Carli, L. L. (1989). Gender differences in interaction style and influence. *Journal of Personality and Social Psychology, 56*(4), 565–576.
de Bruijn, E. R. A., Mars, R. B., Bekkering, H., ... Coles, M. G. H. (2012). Your mistake is my mistake… or is it? Behavioural adjustments following own and observed actions in cooperative and competitive contexts. *The Quarterly Journal of Experimental Psychology, 65*(2), 317–325.
de Bruijn, E. R. A., Miedl, S. F., & Bekkering, H. (2011). How a co-actor's task affects monitoring of own errors: Evidence from a social event-related potential study. *Experimental Brain Research, 211*(3–4), 397–404.
Deakin, J. M., & Proteau, L. (2000). The role of scheduling in learning through observation. *Journal of Motor Behavior, 32*(3), 268–276.
Granados, C., & Wulf, G. (2007). Enhancing motor learning through dyad practice. *Research Quarterly for Exercise and Sport, 78*(3), 197–203.
Guadagnoli, M. A., & Lee, T. D. (2004). Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior, 36*(2), 212–224.
Hamilton, A. F. de C., & Lind, F. (2016). Audience effects: What can they tell us about social neuroscience, theory of mind and autism? *Culture and Brain, 4*(2), 159–177.
Hancock, G. R., Butler, M. S., & Fischman, M. G. (1995). On the problem of two-dimensional error scores: Measures and analyses of accuracy, bias, and consistency. *Journal of Motor Behavior, 27*(3), 241–250.
Harskamp, E., Ding, N., & Suhre, C. (2008). Group composition and its effect on female and male problem-solving in science education. *Educational Research, 50*(4), 307–318.
Intrinsic Motivation Inventory. (n.d.). Retrieved from http://selfdeterminationtheory.org/intrinsic-motivation-inventory/.
Karlinsky, A., Alexander, B., & Hodges, N. J. (in preparation). Partner-related error compensation in a multi-skill learning protocol: You're wrong, I'll switch, I'm wrong, I'll stay.
Karlinsky, A., & Hodges, N. J. (2014). Evaluating the effectiveness of peer-scheduled practice. *Journal of Motor Learning and Development, 2*(4), 63–68.
Karlinsky, A., & Hodges, N. J. (2018a). Dyad practice impacts self-directed practice behaviors and motor learning outcomes in a contextual interference paradigm. *Journal of Motor Behavior, 50*(5), 579–589.
Karlinsky, A., & Hodges, N. J. (2018b). Turn-taking and concurrent dyad practice aid efficiency but not effectiveness of motor learning in a balance-related task. *Journal of Motor Learning and Development, 6*(1), 35–52.
Lee, T. D., Swinnen, S. P., & Serrien, D. J. (1994). Cognitive effort and motor learning. *Quest, 46*(3), 328–344.
Lee, T. D., & White, M. A. (1990). Influence of an unskilled model's practice schedule on observational motor learning. *Human Movement Science, 9*(3–5), 349–367.
Lee, T. D., Wishart, L. R., Cunningham, S., & Carnahan, H. (1997). Modeled timing information during random practice eliminates the contextual interference effect. *Research Quarterly for Exercise and Sport, 68*(1), 100–105.
Lewthwaite, R., & Wulf, G. (2012). Motor learning through a motivational lens. In N. J. Hodges, & A. M. Williams (Eds.). *Skill Acquisition in Sport: Research, Theory and Practice* (pp. 173–191). (2nd ed.). London, UK: Routledge.
McNevin, N. H., Wulf, G., & Carlson, C. (2000). Effects of attentional focus, self-control, and dyad training on motor learning: Implications for physical rehabilitation. *Physical Therapy, 80*(4), 373–385.
Mussi, D. R., Marino, B. F., & Riggio, L. (2015). The influence of social and nonsocial variables on the Simon effect. *Experimental Psychology, 62*(4), 215–231.
Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia, 9*(1), 97–113.
Picton, L., Saunders, B., & Jentzsch, I. (2012). "I will fix only my own mistakes": An ERP study investigating error processing in a joint choice-RT task. *Neuropsychologia, 50*(5), 777–785.
Rhea, M. R., Landers, D. M., Alvar, B. A., & Arent, S. M. (2003). The effect of competition and the presence of an audience on weight lifting performance. *Journal of Strength & Conditioning Research, 17*(2), 303–306.
Rumenik, D. K., Capasso, D. R., & Hendrick, C. (1977). Experimenter sex effects in behavioral research. *Psychological Bulletin, 84*(5), 852–877.
Schmitz, L., Vesper, C., Sebanz, N., & Knoblich, G. (2017). Co-representation of others' task constraints in joint action. *Journal of Experimental Psychology: Human Perception and Performance, 43*(8), 1480–1493.
Shea, C. H., Wright, D. L., Wulf, G., & Whitacre, C. (2000). Physical and observational practice afford unique learning opportunities. *Journal of Motor Behavior, 32*(1), 27–36.
Shea, C. H., Wulf, G., & Whitacre, C. (1999). Enhancing training efficiency and effectiveness through the use of dyad training. *Journal of Motor Behavior, 31*(2), 119–125.
Sherwood, D. E. (2010). Parameter value switching in discrete and continuous aiming movements. *Perceptual & Motor Skills, 111*(3), 901–917.
Sherwood, D. E., & Fosler, J. (2013). Blocked and alternating variable practice and unintended spatial variations in continuous aiming movements. *Perceptual & Motor Skills, 116*(2), 611–625.
Sherwood, D. E., & Rothman, K. K. (2011). Concurrent visual feedback and spatial accuracy in continuous aiming movements. *Perceptual & Motor Skills, 113*(3), 825–839.
Simon, D. A., & Bjork, R. A. (2002). Models of performance in learning multisegment movement tasks: Consequences for acquisition, retention, and judgments of learning. *Journal of Experimental Psychology: Applied, 8*(4), 222–232.
van der Weiden, A., Aarts, H., Prikken, M., & van Haren, N. E. (2016). Individual differences in action co-representation: Not personal distress or subclinical psychotic experiences but sex composition modulates joint action performance. *Experimental Brain Research, 234*(2), 499–510.
van der Wel, R. P., & Fu, E. (2015). Entrainment and task co-representation effects for discrete and continuous action sequences. *Psychonomic Bulletin & Review, 22*(6), 1685–1691.
Wright, D. L., Li, Y., & Coady, W. (1997). Cognitive processes related to contextual interference and observational learning: A replication of Blandin, Proteau, and Alain (1994). *Research Quarterly for Exercise and Sport, 68*(1), 106–109.
Wulf, G., & Lewthwaite, R. (2016). Optimizing performance through intrinsic motivation and attention for learning: The OPTIMAL theory of motor learning. *Psychonomic Bulletin & Review, 23*(5), 1382–1414.
Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review, 9*(2), 185–211.