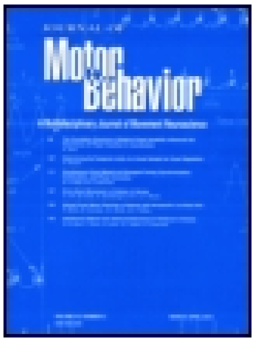

Journal of Motor Behavior

ISSN: 0022-2895 (Print) 1940-1027 (Online) Journal homepage: https://www.tandfonline.com/loi/vjmb20

Learning How to Throw Darts. Effects of Modeling Type and Reflection on Novices' Dart-Throwing Skills

Janneke van der Loo, Emiel Krahmer & Marije van Amelsvoort

To cite this article: Janneke van der Loo, Emiel Krahmer & Marije van Amelsvoort (2020): Learning How to Throw Darts. Effects of Modeling Type and Reflection on Novices' Dart-Throwing Skills, Journal of Motor Behavior, DOI: 10.1080/00222895.2020.1732861

To link to this article: https://doi.org/10.1080/00222895.2020.1732861

Journal of Motor Behavior, 2020

ARTICLE

Learning How to Throw Darts. Effects of Modeling Type and Reflection on Novices' Dart-Throwing Skills

Janneke van der Loo, Emiel Krahmer, Marije van Amelsvoort
Tilburg center for Cognition and Communication, Tilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, Netherlands.

ABSTRACT. In this study, we investigated the effects of modeling type and reflection on the acquisition of dart-throwing skills, self-efficacy beliefs and self-reaction scores by conceptually replicating a study by Kitsantas, Zimmerman, and Cleary (2000). Participants observing a novice model were expected to surpass participants observing an expert model who in turn were expected to outperform participants who learned without a model. Reflection was hypothesized to have a positive effect. 156 High school and university students were tested three times: in a pretest, after a modeling intervention, and after a practice round. Contrary to what was expected, we found no main effects of modeling type and reflection. No interaction effects were found either. There was an effect of testing moment, indicating that participants improved dart-throwing skills, self-efficacy beliefs, and self-reaction scores over time. With these findings, we are not able to replicate Kitsantas et al. From our study, we conclude that observational learning, irrespective of the model's skill level, combined with physical practice, yields similar results as mere physical practice.

Keywords: darts, modeling type, motor skill learning, observational learning, reflection

Lay Summary

Mastering a motor skill, such as dart throwing, is often done by practicing darts, observing someone else playing darts or a combination of observing and practicing. In our study, we find no differences between students who learn by observation combined with physical practice and those who merely practiced.

Introduction

Acquiring a motor skill, such as throwing darts, is generally not learned from reading a book, but rather by observing others playing darts, by practicing the throws yourself or a combination of observational learning and physical practice. In this paper on the comparison of observation and physical practice in learning a motor skill, more specifically throwing darts, we address three issues. First, we aim to determine how observational learning, both with and without physical practice, contributes to learning a motor skill compared to physical practice. Second, we address the concerns related to the skills level of the model, by comparing observational learning from a novice model and an expert model with physical practice. Third, we explore the role of reflection by the learner in both observational learning and physical practice. We will do so by conceptually replicating an earlier study by Kitsantas et al. (2000).

Observation versus Physical Practice

Observational learning is the process of learning a new task by watching someone else performing this task. It relies on multiple capabilities: learners should be able to infer the intentions of others from action observation, process others' action outcomes and combine these sources of information in order to select behaviors leading to desired outcomes later on (e.g., Bandura, 1977; Monfardini et al., 2013; Rak, Bellebaum, & Thoma, 2013). With action observation, the learner does not have to generate a simulated representation of the movement, as the key perceptual information is provided in the form of an external stimulus being observed (Ram, Riggs, Skaling, Landers, & McCullagh, 2007). Arguably, this reduces the complexity of the learning task, since the observer is not as cognitively engaged in the task compared to a performer of that same task. Therefore, the observer can more easily break the whole task in subcomponents, selecting essential information and constructing appropriate strategies to reconstruct the task (Cordovani & Cordovani, 2016; Wulf, Shea, & Lewthwaite, 2010).

Observational learning is generally found to be more effective than doing nothing (e.g., D'Innocenzo, Gonzalez, Williams, & Bishop, 2016) or than verbal instructions alone (Janelle & Hillman, 2003). However, the literature is inconclusive when it comes to comparison of mere practice, mere observation and the combined effects of practice and observation. Some studies imply that observational learning combined with physical practice leads to more effective ways of performing a task

---

Correspondence address: Janneke van der Loo, Tilburg center for Cognition and Communication, Tilburg University, PO Box 90153, 5000 LE, The Netherlands. E-mail: J.M.vdrLoo@tilburguniversity.edu

This article has been republished with minor changes. These changes do not impact the academic content of the article.

than mere physical practice (Horn, Williams, Scott, & Hodges, 2005; Kitsantas et al., 2000; Shea, Wright, Wulf, & Whitacre, 2000; Weeks & Anderson, 2000), while in other studies it is suggested that there is no advantage of observation combined with physical practice over mere physical practice (Weir & Leavitt, 1990). What complicates the comparisons between these studies, is that in the majority of studies (except for Weir & Leavitt, 1990), physical practice is compared to observation combined with physical practice, confounding the modeling effects with the subjects' practice. In the current study, we therefore aim to determine how observational learning and physical practice contribute to learning a motor skill by systematically comparing observational learning with and without physical practice to mere physical practice.

Type of Model in Observational Learning

Within observational learning the skill level of the model may influence the learning process (Kitsantas et al., 2000; Shea et al., 2000; Wesch, Law, & Hall, 2007). Some have advocated the use of expert models, since they display a standard of reference against which observers are able to detect their own errors and issue appropriate correction, which facilitates constructing a mental representation (Carroll & Bandura, 1987; Ferrari, 1996). Others have advocated the use of novice models, gradually improving their performance, since they provide more information about strategy implementation and error correction than expert models do (Darden, 1997; Ferrari, 1996; Zimmerman & Kitsantas, 2002).

Studies explicitly comparing the effects of novice and expert models in motor skill acquisition have yielded equivocal results. Kitsantas et al. (2000) found that observing a novice model leads to more effective learning of a motor skill than observing an expert model, which in turn leads to better learning than learning without a model, while in a study by Rohbanfard and Proteau (2011) observing an expert model leads to more stable performances with less variability than observing a novice model. Several other studies, on the other hand, found no effects of modeling type on different types of motor skill performance (see e.g., Blanchard, 2014; Blandin, Lhuisset, & Proteau, 1999; Moore, Lelievre, & Ste-Marie, 2019; Pollock & Lee, 1992; Weir & Leavitt, 1990). In these latter studies, physical practice or combinations of physical practice and observation, irrespective of the models' skill level, increased performance.

In view of these contrasting findings, we aim to further explore which model most effectively promotes learning a motor skill by systematically comparing the observation of a novice model, an expert model and observing no model in learning how to throw darts.

Reflection in Observational Learning and Physical Practice

The effects of observational learning are often explained by the fact that it allows learners to construct a mental representation of the desired performance. The observer watches a model and transfers the provided information to his or her own acquisition by judging which parts could be beneficial and how they could be used. It is hereby of importance that observers are capable of identifying relevant features or key movement pattern elements of a motor skill. Adding instructional components to the modeled information could facilitate identifying these relevant cues. However, only a limited number of studies have included such instructional components directing attention to these relevant cues in motor skill learning (Ste-Marie et al., 2012).

In more cognitive research domains, mostly within writing research in which students learn how to write (fairly) complex texts, instructional components are included in the observational learning process. Observers are encouraged to carry out different cognitive, reflective activities, by asking them for example to monitor, evaluate and elaborate on models' performances (e.g., Braaksma, Rijlaarsdam, Van den Bergh, & Hout-Wolters, 2006; Raedts, Van Steendam, De Grez, Hendrickx, & Masui, 2017). These reflective activities support students in developing criteria for effective texts and writing processes which transfer to their own writing, yielding higher quality writing performance, in terms of the degree of selection of relevant and correct information, level and quality of integration of selected ideas, and textual organization (Braaksma, van den Bergh, Rijlaarsdam, & Couzijn, 2001).

From these findings, it can be concluded that reflection arguably supports learners in developing a mental representation of the desired performance and in developing performance strategies which might lead to a more successful performance. In the current study, we therefore explore whether these findings from writing research can be transposed to motor skill learning, by adding reflection to the design.

Aim of the Study

In the current study, we investigate the effect of modeling type (novice, expert and no model) and reflection on acquiring a motor skill. Since previous studies on the effect of modeling type in learning a motor skill vary substantially in their designs and analyses, we chose to conceptually replicate one of the earlier studies, Kitsantas et al. (2000), which allows us to test the robustness of the theoretical implications (Zwaan, Etz, Lucas, & Donnellan, 2018). We chose this particular study because from the handful of studies directly comparing effects of novice and expert models on motor

TABLE 1. Dependent variables measured in the tests.

	Dart throwing	Self-efficacy	Self-reaction	Error attribution	Dart technique*
Pretest	+	+	+		
Post-test	+	+	+		+
Delayed post-test	+	+	+	+	

*Dart Technique was only measured for the high school students.

learning, this one resulted in significant differences in the effectiveness of the different types of models, and is conceptually closest to what we aim to explore.

Kitsantas et al. (2000) investigated the effect of modeling type on dart throwing performance and self-regulatory measures with 60 female high school students. The students were first presented with written instructions, followed by either watching a video of a novice model or an expert model, or not watching any video (physical practice condition). The girls were then given fifteen minutes to practice dart throwing, after which they were tested on dart-throwing performance, self-efficacy, error attribution, self-reactions, and intrinsic interest. Kitsantas et al. (2000) found that the girls who had observed the novice model outperformed the girls who had observed the expert model on all measures. The girls who observed an expert model scored in turn better than the girls in the physical practice condition.

In the current study, we closely follow the procedures and measures of Kitsantas et al. (2000) which will be described in the method section. However, we extended the original design in four aspects. First, we add reflection as a factor to the design. In Kitsantas et al. (2000), some reflective activities were present in the novice model condition, but absent in the expert model condition and in the physical practice condition, which complicates comparing the different conditions. Second, we extended our sample to not only include high school students but also university students. Third, we add a measure, implementation of dart techniques, to establish whether modeling type influences the physical execution of the performance. Finally, we test the participants three times during the experiment, while participants in the original study were only tested once. This allows us to test possible direct effects of observational learning and combined effects of observation and physical practice.

Because of the replication character of the current study, our hypotheses are similar to Kitsantas et al. (2000). First of all, modeling type is expected to influence the acquisition of a novice motor skill. We hypothesize that (H1a) observing a model leads to a higher performance on all measures than not observing a model, and that (H1b) observing a novice model leads to a higher performance than observing an expert model. Secondly, based on earlier findings concerning the influence of reflection in writing research (Braaksma et al., 2001), we hypothesize that (H2) reflection leads to a higher performance on all measures than no reflection. And thirdly, by adding reflection to physical practice, an interaction effect between modeling type and reflection is hypothesized, in which we expect that (H3) reflection reduces the effect of modeling type on all measures.

Method

Design

A 3 × 2 design was used in this experiment, with modeling type (novice model, expert model and no model) and reflection (yes or no) as the between-subjects factors. This resulted in six conditions to which 156 participants were randomly assigned. The experiment was run with university students and with high school students, in order to increase power and generalizability. In Kitsantas et al. (2000) the sample size was 60. We therefore aimed at a minimum of 150 participants, based on the suggestion by Simonsohn (2015) that a way to determine sample sizes in replication studies is to take 2.5 times the original sample size.

The participants were tested on dart-throwing skills, self-efficacy, self-reaction and error attribution. We tested dart-throwing skills, self-efficacy and self-reaction three times within the experiment. The first test (pretest) was at the start of the experiment, so we could establish a baseline. The second test (post-test) was directly after the experimental treatment, in order to measure direct effects of observational learning from a novice model, an expert model, or of physical practice. The third test (delayed post-test) took place after a ten-minute practice session, in order to determine combined effects of observational learning and physical practice. Error attribution was only measured during the delayed post-test. In the experiment with the high school students we added implementation of the dart techniques as a measure, in order to detect possible differences in the way the darts were thrown. This was measured during the post-test. Table 1 displays which dependent variables were measured in each test.

TABLE 2. Distribution of participants over conditions.

	Reflection		No Reflection		Total
	University	High school	University	High school	
Novice Model	16	11	15	11	53
Expert Model	15	11	15	11	52
No Model	14	11	15	11	51
Total	78		78		156

Participants

In this study 156 students (88 women, equally distributed over conditions) participated. Ninety of these participants were undergraduate students recruited from a participant pool at a Dutch University (*M_age* = 21,2 years, *SD* = 3.87). The other 66 participants were Dutch high school students (*M_age* = 14,3 years, *SD* = 0.88) who took part in the experiment voluntarily. Table 2 presents an overview of the distribution of the participants over conditions.

Materials

Dart Board

Identical to the study by Kitsantas et al. (2000), we used a dart board with regular concentric circles. The dart board had a bullseye and nine consecutive circles, with each circle having a width of 2.54 centimeter. A numerical value was assigned to each circle, beginning with a value of 10[1] for the bullseye in the center and successively diminishing in assigned values by 1 until the outermost circle, which had a value of 1. The dart board was positioned, in line with the rules of the British Darts Association (2016), at a distance of 2.37 meters of the throwing line, with the bullseye at a height of 1.73 meters from the ground. The participants were provided with six regular darts so they could make six consecutive throws during the tests.

Hand-out Dart Instructions

The participants were provided with a hand-out which contained instructions for dart throwing. The instructions included information on five subskills: grip, stance, sighting, throw and follow through, and were derived from the instructions used in Kitsantas et al. (2000) and translated into Dutch. Table 3 provides an overview of the instructions.

Videos

Two different videos were recorded, one of a novice model playing darts (see Figure 1) and one of a expert model playing darts. The same male model appeared in both videos. The videos displayed information on all five subskills described on the hand-out and in each of them the model threw 15 darts. In the novice model condition, the model started out by making several errors in his dart-throwing technique. These errors were made by not following the instructions to dart-throwing that were described, for instance by holding the dart with all five fingers instead of the three mentioned in the instruction. The model commented on his own performance and gradually improved his dart-throwing skills. In the expert model video, the model started out with a flawless technique, and maintained this throughout the video. The model commented on his own performance, by describing the techniques he was using. Both videos lasted approximately 1 minute and 40 seconds.

Reflection Questions

Participants in the reflection conditions were asked to answer reflection questions on each of the five subskills from the instructions. Given the differences in content between the conditions, reflection questions also differed somewhat per condition. In the expert-model with reflection condition, participants were asked open-ended questions for each subskill. An example for the subskill *stance* is “How did the dart-thrower position his feet?”. In the novice-model with reflection condition, participants were asked to answer additional questions for each subskill, namely “What kind of mistakes did the model make?” and “How did the model correct his mistakes?”. In the no-model with reflection condition, participants reflected on their personal performance, for example “How did you position your feet?”. They were then asked what kind of mistakes they made (if any), and how they would correct these the next time. These questions allowed participants to reflect on and process effective and/or less effective strategies in all conditions.

Measures

Dart-Throwing Skills

To measure dart-throwing skills, we added up the points for all six consecutive throws. The minimum score per dart was zero and the maximum score was ten. The average dart-throwing score per dart was calculated

TABLE 3. Dart-throwing instructions provided to the participants (taken from Kitsantas et al., 2000).

Subskill	Instruction
Grip	Hold the dart between your first and second finger and the thumb. Simply grasp the dart comfortably.
Stance	Stand behind the white throwing line facing the target. Stand comfortably with your feet slightly apart. If you are right-handed, the right foot should be slightly ahead of the left, touching the toe line and pointing toward the board. If you are left-handed, place the left foot forward.
Sighting	Keep your arm close to your body. Using your arm and wrist, and with the elbow acting as a fulcrum, bring the dart back toward your face it until almost brushes your cheek where you find it most comfortable to stop.
Throw	Keep all the other parts of your body still when you throw. Your head must be held steady, and you must not jerk the throw. Try to develop a smooth, vertical throw using the wrist and elbow as pivots. Hold your elbow steady, and keep it parallel to the floor. Your wrist should be loose and laid slightly back. Use only the forearm and wrist to throw in a vertical motion. The throw need not be hard, but it must be crisp. The dart should get to the board quickly with as little trajectory as possible.
Follow-through	After you release the dart, simply allow your arm to continue in its natural motion. Let your hand, with your fingers fully extended, follow the dart as it moves toward the target.

Note. The instructions were translated into Dutch before providing them to the participants.

FIGURE 1. Still of the model playing darts.

by dividing the total score by six. Dart-throwing skill was measured in all three tests.

Self-Efficacy

Before each dart-throwing test, participants were asked how confident they felt that they could (1) throw 9 points with one dart, (2) throw 7 points with one dart, (3) throw 5 points with one dart, and (4) throw 3 points with one dart. Their scores were measured on a scale from 0 to 100 with 10-point intervals, with 0 being not sure and 100 being very sure. This self-efficacy measure was comparable to the measure used by Kitsantas et al. (2000). For each of the three tests, the average score of these four questions was used as the participants' self-efficacy score (Cronbach's Alpha in all three tests $> .86$).

Self-Reaction

Participants had to indicate for each of the three tests on a scale from 0 to 100 how satisfied they were with their own performance, with 0 being not satisfied and 100 being very satisfied. This measurement was identical to that of Kitsantas et al. (2000).

Error Attributions

After the delayed post-test, we asked the participants why they thought they had missed the bullseye. If they hit the bullseye, they could skip this question: this was the case for five participants. Their answers were grouped into six categories, namely: type of strategy, amount of effort, level of ability, amount of practice, I don't know, and other. This measurement was only carried out after the last test, in line with Kitsantas et al. (2000).

Dart Techniques

In the experiment with the high-school students, we added the measure dart techniques. In order to measure whether the participants implemented the correct dart technique, we filmed the dart-throwing test of the high-school students' post-test, which followed the experimental treatment immediately. We scored the first and last throw on the five subskills presented in the instructions: grip, stance, sighting, throw and follow through. We scored 1 point if the subskill was implemented correctly, and 0 if it was not implemented correctly. This resulted in a possible minimum score of 0 and a maximum score of 10 (5 for the first throw and 5 for the last throw).

Procedure

The university students were individually taken into a room where they were introduced to the procedure by an experimenter and filled out a consent form. The high school students were taken into a large physical education room in pairs. In the room, two experimenters were available, one for each student. The high school students were each taken to opposite sides of the room and introduced to the procedure individually by one of the two experimenters. Parental consent was obtained by the managing director of the high school, prior to the experiment. During the experiment the high school students did not face each other, since we set up the room in such a way that the two dart boards were on opposite walls of the room. This way the students had their backs turned to each other, with at least 3-5 meters between them. All participants then filled out a form with demographic information, and they performed a baseline test in which their self-efficacy, dart-throwing skills, and self-reaction were measured (pretest). Within this test, participants first answered the question about self-efficacy, then they threw six consecutive darts, and after throwing the darts, they answered the self-reaction question.

After the pretest, they were provided with the handout containing the instructions to dart throwing (grip, stance, sighting, throw, and follow through). The participants were asked to study these until they felt confident that they knew what to do. After having read the instruction, the procedure depended on the condition participants were assigned to.

No Model, No Reflection

Participants were tested again straight after reading the instructions (post-test). Then, they got to practice dart throwing for ten minutes. After the ten-minute practice, the participants were tested for one last time (delayed post-test).

Novice Model, No Reflection

Participants were shown a video of the model who started out by making certain mistakes, but significantly improved during the video. After having watched the video, the participants performed the same test again, thus, their self-efficacy, dart-throwing skills, and self-reaction were measured (post-test). Then, they got to practice dart throwing for ten minutes. After the ten-minute practice, the participants were tested for one last time (delayed post-test).

Expert Model, No Reflection

Participants were shown a video of a dart-thrower who made 15 good throws. After having watched the video, the participants performed the same test again, thus, their self-efficacy, dart-throwing skills, and self-reaction were measured. Then, they got to practice dart throwing for ten minutes. After the ten-minute practice, the participants were tested for one last time (delayed post-test).

No Model, with Reflection

Participants were asked to perform their first fifteen practice throws. After that, they were asked to reflect upon their own performance according to the five subskills described in the instructions (grip, stance, sighting, throw, and follow through). After the reflection, they were tested again on self-efficacy, dart-throwing skills, and self-reaction (post-test). Afterwards, they continued to practice. Because we wanted to make sure that the total practice time was equal across all conditions, the duration of the first fifteen throws of the participants in the no model, with reflection group, was part of their total ten minutes of practice. Thus, if participants performed their first fifteen throws in three minutes, they had an additional seven minutes to practice after the reflection. After the practice period, the participants were tested for one last time (delayed post-test).

Novice Model, with Reflection

After having read the instructions, participants were shown the video of a dart-thrower who started out by making certain mistakes, but significantly improved over the rounds, the model threw 15 darts in total. Participants were asked to evaluate the model's performance on the five subskills explained in the instruction and to report mistakes and possible improvements. After having watched the video and having reflected on the model's performance, the participants performed the same test again, thus, their self-efficacy, dart-throwing skills, and self-reaction were measured. Then, they got to practice dart throwing for ten minutes. After the ten-minute practice, the participants were tested for one last time (delayed post-test).

Expert Model, with Reflection

After having read the instructions, participants were shown the video of a dart-thrower who made 15 good throws. They were asked to evaluate the performance of

TABLE 4. Average Dart-throwing Skills (with SDs) as a function of Reflection (Reflection, No reflection), Model (No Model, Expert, Novice) and Test (pre-, post-, delayed post-test).

	No Model			Expert			Novice		
Test	Pre	Post	Delayed	Pre	Post	Delayed	Pre	Post	Delayed
No reflection	4.0 (1.8)	4.2 (2.0)	5.8 (1.5)	4.7 (1.8)	4.8(1.6)	5.9(1.6)	5.0(1.6)	4.7(1.7)	5.6 (1.5)
Reflection	3.9 (1.4)	4.3 (1.5)	5.4 (1.6)	4.8 (1.4)	5.2 (1.7)	5.9 (1.3)	4.4 (2.2)	4.3 (1.9)	5.6 (1.7)

the model on the subskills explained in the instruction. After having watched the video and having reflected on the model's performance, the participants performed the same test again, thus, their self-efficacy, dart-throwing skills, and self-reaction were measured. Then, they got to practice dart throwing for ten minutes. After the ten-minute practice, the participants were tested for one last time (delayed post-test).

Analyses

A repeated measure multivariate ANOVA was conducted, for each dependent variable (dart throwing skills, self-efficacy and self-reaction), with modeling type (no model, expert model, novice model), reflection (absent, present) and group (university students, high school students) as between subjects-variables and test moment (pre-, post- and delayed post-test) as a within-subject variable. Mauchly's test for sphericity was used to test for homogeneity of variance. We applied Bonferroni corrections for the pairwise comparisons. For attribution, we performed a chi square analysis, since attribution was measured on a categorical score. An ANOVA with modeling type, reflection and groups as independent variables was conducted for measuring implementation of dart techniques.

Results

Dart-Throwing Skills

A within-subject main effect of testing moment was found for dart-throwing skills (F (2,144) = 56.01, $p < .01$, $\eta p^2 = .28$), with the mean scores gradually increasing each testing moment ($M_{pre} = 4.40$, CI = [4.13, 4.68], $M_{post} = 4.47$, CI = [4.21, 4.74], $M_{delay} = 5.62$, CI = [5.40, 5.85]). The mean difference between the pretest and the delayed post-test (*M-diff* = −1.22) was significant, $p < .01$, CI = [−1.48, −.96] and so was the difference between the post-test and the delayed post-test (*M-diff* = −1.15), $p < .01$, *CI* [−1.40, −.90]). However, the difference between the pretest and post-test (*M-diff* = −.07) turned out not to be significant, $p = .59$, CI = [−.33, .19].

Group had a significant main effect on dart-throwing skill, F (1,144) = 16.59, $p < .01$, $\eta p^2 = .10$, with the university students ($M = 5.27$) scoring significantly higher than the high school students ($M = 4.40$).

No significant main effect of modeling type on dart-throwing skills was found, F (2,144) = 2.13, $p = .10$. The data numerically suggest that best results are obtained by observing an expert model ($M = 5.10$, CI = [4.73, 5.47]), followed by a novice model (($M = 4.86$, CI = [4.51, 5.22]) and no model ($M = 4.54$, CI = [4.17, 4.90]).

Reflection had no main effect on dart-throwing skill either, F (1,144) = 0.08, $p = .77$. The dart-throwing scores in the conditions without reflection ($M = 4.86$, CI = [4.57, 5.16]) were not significantly different from those in the conditions with reflection ($M = 4.80$, CI = [4.51, 5.10]).

In addition, none of the interaction effects were found to be significant ($F_s < 1.43$, $p_s > .24$). Table 4 displays the mean dart-throwing skills scores per condition.

Self-Efficacy

The pattern for self-efficacy is very comparable to that of dart-throwing skills. Again, a main effect of testing moment on self-efficacy was found, F (1.77, 254.48) = 185.40, $p < .01$, $\eta p^2 = .56$. Self-efficacy scores on the pretest (M = 49.93, CI = [47.22, 52.63]), the post-test (M = 59.78, CI [57.17, 62.39]), and the delayed post-test (M = 68.56, CI = [66.01, 71.11]) were significantly different. All the individual differences were significantly different at a p-value of < .01.

Group also had a significant main effect on self-efficacy, F (1,144) = 49.68, $p < .01$, $np^2 = .26$, with the university students ($M = 50.94$) scoring significantly lower than the high school students ($M = 67.90$).

The main effect of modeling type did not yield significance, F (2,144) = 2.73, $p = .07$. Again, the data numerically suggest that best results are obtained by observing an expert model (M = 63.42, CI [59.25, 67.59]), followed by observing a novice model (M = 57.64, CI [53.60, 61.68]) and no model (M = 57.21, CI [53.06, 61.35]). However, this effect failed to reach significance.

The main effect of reflection was not found to be significant either, F (1,144) = 3.62, $p = .06$. Similar to the

TABLE 5. Average Self-Efficacy (with SDs) as a function of Reflection (Reflection, No reflection), Model (No Model, Expert, Novice) and Test (pre-, post-, delayed post-test).

	No Model			Expert			Novice		
Test	Pre	Post	Delayed	Pre	Post	Delayed	Pre	Post	Delayed
No reflection	50.6 (20.0)	56.7 (18.2)	67.3 (18.2)	51.5 (16.9)	66.0 (15.7)	75.1 (13.7)	49.5 (19.6)	60.0 (17.8)	66.7 (16.2)
Reflection	43.5 (21.3)	52.8 (19.3)	64.0 (18.5)	49.6 (17.7)	60.4 (17.4)	68.9 (65.1)	44.7 (23.4)	55.3 (19.6)	63.5 (18.0)

TABLE 6. Average Self-reaction Scores (with SDs) as a function of Reflection (Reflection, No reflection), Model (No Model, Expert, Novice) and Test (pre-, post-, delayed post-test).

	No Model			Expert			Novice		
Test	Pre	Post	Delayed	Pre	Post	Delayed	Pre	Post	Delayed
No reflection	63.9 (17.0)	56,92 (25,082)	71.1 (19.9)	66.5 (19.8)	58.6 (14.1)	73.2 (15.6)	65.5 (17.7)	56.2 (19.7)	63.3 (18.9)
Reflection	54.8 (21.9)	52.5 (22.3)	60.5 (15.1)	66.8 (13.9)	62.5 (23.1)	67.8 (21.9)	64.9 (18.5)	54.4 (22.4)	66.6 (22.0)

effect of reflection on dart-throwing skills, the data numerically suggest that highest scores are obtained by not reflecting (M = 61.71, CI = [58.35, 65.07]) instead of reflecting (*M* = 57.13, CI = [53.77, 60.50]). Again, no significant interaction effects were found ($F_s < 0.59$, $p_s > .55$). In Table 5, the mean self-efficacy scores per condition are displayed.

Self-Reaction

The pattern of results for self-reaction was very similar to those of self-efficacy and dart-throwing skills. A significant main effect of testing moment on self-reaction scores was found, F (2, 144) = 14.93, $p < .01$, $\eta p^2 = .09$. In the post-test the scores dropped compared to the pretest (M_{pre} = 63.23, *C.I.* [60.31, 66.15], M_{post} = *56.39,* CI = [52.94, 59.84]. In the delayed post-test (M_{delay} = 67.14, CI = [64.07, 70.21]) the mean scores increased again. The mean difference between the pretest and 2 (*M-diff* = 6.84) was significant, $p < .01$, CI = [1.72, 11.96] and so was the mean difference between the post-test and delayed post-test (*M-diff* = −10.74), $p < .01$, *CI* [−15.68, −5.81]). However, the mean difference between the pretest and delayed post-test (*M-diff* = −3.91) turned out not to be significant, $p = .10$, CI = [−.8.28, .47].

No significant main effect of modeling type, (F (2,144) = 2.44, p = .09) or reflection (F (1,144) = 1.54, p = .22) was found, indicating that no model (M = 59.85, CI [56.04, 63.67]), observing an expert model (M = 65.65, CI [61.81, 69.50]) and observing a novice model (M = 61.25, CI [57.53, 64.97]) do not lead to significantly different self-reaction scores, and neither do not reflecting (M = 63.63, CI [60.53, 66.72]) and reflecting (M = 60.88, CI [57.78, 63.97]). Again, the numerical differences suggest best results are obtained by observing an expert model, followed by a novice model and learning by doing, and by not reflecting instead of reflecting.

Group had no effect on self-reaction, F (1,144) = 2.61, p = .11 In addition, no significant interaction effects were found, $F_s < 1.54$, $p_s > .22$. See Table 6 for an overview of the self-reaction scores.

Attribution

Table 7 shows the frequency of each answer for the different modeling conditions, whereas these frequencies are displayed for the reflection conditions in Table 8.

Before the chi-square analyses could be performed, we recoded the answers into four categories, combining Effort, Other and Don't Know in order for all cells to have a count above five, which is required for a chi-square analysis. The analysis showed no significant differences between the modeling groups, $\propto (6) = 2.75$, $p = .84$ and between the reflection groups, $\propto (3) = 2.60$, $p = .46$. These results indicate that there is no significant difference between participants who learned from observing an expert model, a novice model or no model, and between those who reflected and those who did not in what they attribute their errors to.

Implementation of Dart Techniques

In the experiment with the high-school students, we measured whether the participants implemented the dart

TABLE 7. Attribution (Count and Percentage) as a function of modeling type.

	Attributional source					
Group	Strategy	Effort	Ability	Practice	Don't know	Other
Control	6 (12%)	3 (6%)	20 (39%)	17 (33%)	3 (6%)	2 (4%)
Expert	7 (15%)	3 (6%)	19 (39%)	14 (29%)	1 (2%)	4 (8%)
Novice	9 (17%)	3 (6%)	25 (48%)	11 (21%)	1 (2%)	3 (6%)

TABLE 8. Attribution (count and percentage) as a function of reflection group.

	Attributional source					
Group	Strategy	Effort	Ability	Practice	Don't know	Other
No Reflection	10 (14%)	2 (3%)	29 (39%)	25 (34%)	5 (7%)	3 (4%)
Reflection	12 (15%)	7 (9%)	35 (46%)	17 (22%)	0 (0%)	6 (8%)

TABLE 9. Scores on Dart Techniques Implementation (SD) as a function of modeling type and reflection (min. score 0, max. score 10).

	No Model	Expert Model	Novice Model	Total
Reflection	6.9 (1.8)	7.3 (1.7)	6.8 (1.9)	7.0 (1.9)
No Reflection	6.5 (1.8)	7.3 (2.1)	7.6 (1.6)	7.1 (1.7)
Total	6.7 (1.8)	7.3 (1.8)	7.2 (1.8)	

techniques in a correct manner. The ANOVA revealed no significant effects of modeling type, *F* (2,53) = 0.64, *p* = .53 or reflection, *F* (1,53) = 0.07, *p* = .80. No interaction was found either, *F* (2,53) = 0.57, *p* = .57. These results indicate that modeling type and reflection did not influence how participants used the techniques when throwing darts. An overview of the mean scores is displayed in Table 9.

Correlations

To further explore the relationships between dart-throwing skills, self-efficacy and self-reaction, we used a Pearson's correlation test. In the pretest, the results show a weak but significant correlation between self-efficacy and dart-throwing skills, $r = .29$, n = 156, $p < .01$., and a medium strong correlation between self-reaction and dart-throwing skills, $r = .47$, n = 156, $p < .01$. The correlation between self-efficacy and self-reaction, however, was very weak and non-significant, $r = .05$, n = 156, $p = .52$.

For the post-test, the correlations are very similar to the pretest. Again, the correlation between self-efficacy and dart-throwing skills is weak but significant, $r = .24$, n = 156, $p < .01$, and the correlation between self-reaction and dart-throwing skills is of medium strength and significant, $r = .51$, $p < .01$. The correlation between self-efficacy and self-reaction is again not significant and weak in this second test, $r = .05$, n = 156, $p = .54$.

In the delayed post-test, again we find a weak but significant correlation between self-efficacy and dart-throwing skills, $r = .28$, n = 156, $p < .01$, and a medium strong correlation between self-reaction and dart-throwing skills, $r = .57$, $p < .01$. However, in the delayed post-test, we do find a weak but significant correlation between self-efficacy and self-reaction, $r = .22$, n = 156, $p < .01$.

General Discussion

Observational learning is often included in learning a complex motor skill, such as dart throwing. By action observation of a model performing the motor skill, learners are arguably less cognitively involved in the task, which could support them in constructing a mental representation of the action, leading to a higher or more accurate performance. Since the findings in previous research are inconclusive on the effects of different modeling types, in this study we sought to determine what would be the most effective way of learning a motor skill by systematically comparing observational learning from a novice model, observational learning from an expert model and learning from physical practice. In addition, we explored whether adding reflective activities to both observational learning and physical practice could enhance learning.

Modeling Type: Expert Model and Novice Model Versus Physical Practice

We found that observational learning from a novice model, an expert model and learning from physical practice yielded similar results on dart throwing skills, self-efficacy beliefs, self-reaction, error attribution, and implementation of dart techniques, which means we have not found support for H1a an H1b. This was the case in both the post-test, directly following the modeling intervention, and the delayed post-test, following the practice session[2].

In the post-test, participants who watched one of the models, had similar darts scores as those who practiced, and they also reported similar levels of self-efficacy. It should be noted, however, that dart throwing performance did not change significantly from the pretest to the post-test, so it could also be argued that a short practice period or mere observational learning does not suffice for learning to throw darts. Another interesting find is that in all groups satisfaction dropped significantly from the pretest to the post-test, while the dart scores did not change. The participants were thus less satisfied with the same score in the post-test, which implies that the learners thought they had learned something from either watching the model or the short practice period. In the post-test, we also measured how accurate the high school students implemented the dart techniques presented to them. Again, no differences were found between conditions. Observational learning did not lead to a more accurate dart throwing technique. A limitation is that we only measured implementation of dart techniques in this post-test. This did allow us to determine immediate effects of observational learning compared to physical practice. However, this prevents us from making statements on the combined effects of observational learning on the implementation of dart techniques.

The results of the delayed post-test, again reveal no differences between observational learning combined with physical practice and mere physical practice. After a ten-minute practice session, all groups significantly improved their dart throwing performance. In addition, all groups showed more confidence in themselves and were more satisfied with their performance, compared to the pretest and post-test, irrespective of condition. This was also the case for error attribution.

In the current study we were not able to replicate the findings by Kitsantas et al. (2000) who found clear effects of modeling type, even though our experimental procedures and measurements were very similar. Our test was also sufficiently powered: our sample was larger than in Kitsantas et al. (respectively 156 and 60) and the test was sensitive enough to find statistical differences for testing moment, indicating that the participants did learn from the interventions. Furthermore, we also found no indications, contrary to Horn et al. (2005), that observational learning facilitates adopting motion information, leading to a more accurate performance.

Our results are, however, in line with Blanchard (2014), Blandin et al. (1999), Pollock and Lee (1992), and Weir and Leavitt (1990) who also found no differences between modeling types in learning a motor skill. The volatile results on the effects of modeling type on motor skill acquisition frustrate drawing clear conclusions on the implications of our study. Our study and these latter studies seem to indicate that observational learning, of either a novice or expert model, does not positively affect learning compared to physical practice, but it also seems not to hinder learning a motor skill. This seems to be especially the case when it is combined with a physical practice period in which observers can experience sensory feedback which arguably improves muscle control (Blandin, Proteau, & Alain, 1994; Cordovani & Cordovani, 2016) leading to a higher performance.

A limitation to our study is arguably the way we measured dart performance. Since our study was a replication of Kitsantas et al. (2000), we measured dart performance in the same way as they did, to facilitate comparison of the results. However, other researchers, including Hancock, Butler, and Fischman (1995) and Fischman, et al. (2015) have argued that with this type of assessment of dart performance is not optimal, because it does not allow determining performance variability, which is considered to be an important characteristic of motor skill learning. Using two-dimensional error measures could support quantifying accuracy, bias and consistency of performance. Future studies should consider adding these two-dimensional error measures to the design. Another limitation is that we did not include a mixed observation group in the design. There are some indications that observing both a novel and expert model might lead to a more stable performance (Rohbanfard & Proteau, 2011), which could be explored in future studies.

Reflection

Previous studies have suggested that integrating instructional components to a motor skill learning regiment could enhance learning, since these components focus attention to relevant movements (e.g., Ste-Marie et al., 2012). In our study, we systematically compared reflecting to not reflecting in both observational learning with a novice and expert model and in physical practice.

In the current study, reflection did not affect dart throwing skills, self-efficacy beliefs, self-reaction, error attribution, and implementation of dart techniques, which means H2 and H3 are not supported by our results. Participants who reflected did not outperform those who did not[3]. It is worth noting that reflection in the different learning conditions was of necessity subtly different: in observational learning the students reflected on the model's performance while in the physical practice condition, they reflected on their own performance, and this difference might conceivably influence the effectiveness of reflection.

Since we found no effects of reflection, it could be argued that the positive effects of reflection found in observational learning within cognitive domains, such as writing, do not transpose to the motor skill domain. The reflection questions did direct attention to relevant features of the performance, but only after the participants' own performance or that of the models. This means that they had to reflect on all five subskills in hindsight, while compared to the writing domain, the subskills are executed within a very limited time span. This might have affected the effectiveness of the reflection questions.

Conclusion

In the current study, we investigated how observational learning from a novice model or an expert model influence dart throwing skills, self-efficacy beliefs, self-reaction, error attribution and implementation of dart techniques, compared to learning from physical practice. In addition, we sought to explore how reflection affects acquiring a motor skill, and what the interplay is between reflection and observational learning from a novice model, an expert model or physical practice. In our study, it did not matter how someone was taught how to play darts, as long as they practiced, their scores improved, they felt more confident and they were more satisfied. Our study contributes to the growing body of research on observational learning within the motor skill domain. From these studies, including the current one, no clear view can be established on the effects of observational learning of either a novice or expert model, whether it is combined with physical practice or not, compared to mere physical practice.

Notes

1. In Kitsantas et al., (2000) the dart board consisted of seven concentric circles, which resulted in a maximum average dart-throwing score of 7.
2. To quantify the evidence of the absence of the effect, we performed post-hoc equivalence tests (Lakens, 2017) comparing the modeling conditions for all dependent variables. The equivalence tests show that the largest equivalence bounds (based on Cohen's D) for dart performance, self-efficacy, self-reaction and implementation of dart techniques vary between .48 and .78, meaning that with current sample size, we can similarly reject medium to large and large effects.
3. To quantify the evidence of the absence of the effect, we performed post-hoc equivalence tests (Lakens, 2017) comparing the reflection conditions for all dependent variables. The equivalence tests show that the largest equivalence bounds (based on Cohen's D) for dart performance, self-efficacy, self-reaction and implementation of dart techniques vary between .34 and .53, indicating that with our sample size, we can reject medium and large effects.

ACKNOWLEDGMENTS

We thank Eefje Frissen for her contributions to this study, as well as Stefan van Strien for his assistance with the data collection.

REFERENCES

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Blanchard, A. (2014). *Bull's eye-hand coordination: Visual and motor contributions to observational learning* (Doctoral dissertation). Rutgers University-Graduate School-Newark.

Blandin, Y., Lhuisset, L., & Proteau, L. (1999). Cognitive processes underlying observational learning of motor skills. *The Quarterly Journal of Experimental Psychology Section A*, *52*(4), 957–979. doi:10.1080/713755856

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. doi:10.1080/00222895.1994.9941657

Braaksma, M. A., van den Bergh, H., Rijlaarsdam, G., & Couzijn, M. (2001). Effective learning activities in observation tasks when learning to write and read argumentative texts. *European Journal of Psychology of Education*, *16*(1), 33–48. doi:10.1007/BF03172993

Braaksma, M., Rijlaarsdam, G., Van den Bergh, H., & Hout-Wolters, V. (2006). What observational learning in writing courses entails: A multiple case study. *L1 Educational Studies in Language and Literature*, *6*(1), 31–62. doi:10.17239/L1ESLL-2006.06.01.05

Carroll, W. R., & Bandura, A. (1987). Translating cognition into action: The role of visual guidance in observational learning. *Journal of Motor Behavior*, *19*(3), 385–398. doi:10.1080/00222895.1987.10735419

Cordovani, L., & Cordovani, D. (2016). A literature review on observational learning for medical motor skills and anesthesia teaching. *Advances in Health Sciences Education*, *21*(5), 1113–1121. doi:10.1007/s10459-015-9646-5

Darden, G. F. (1997). Demonstrating motor skills—rethinking that expert demonstration. *Journal of Physical Education, Recreation & Dance*, *68*(6), 31–35. doi:10.1080/07303084.1997.10604962

D'Innocenzo, G., Gonzalez, C. C., Williams, A. M., & Bishop, D. T. (2016). Looking to learn: The effects of visual guidance on observational learning of the golf swing. *PloS One*, *11*(5), e0155442. doi:10.1371/journal.pone.0155442

Ferrari, M. (1996). Observing the observer: Self-regulation in the observational learning of motor skills. *Developmental Review*, *16*(2), 203–240. doi:10.1006/drev.1996.0008

Fischman, M. G. (2015). On the continuing problem of inappropriate learning measures: Comment on Wulf et al. (2014) and Wulf et al. (2015). *Human Movement Science*, *42*, 225–231. doi:10.1016/j.humov.2015.05.011

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. doi:10.1080/00222895.1995.9941714

Horn, R. R., Williams, A. M., Scott, M. A., & Hodges, N. J. (2005). Visual search and coordination changes in response to video and point-light demonstrations without KR. *Journal of Motor Behavior*, *37*(4), 265–274.

Janelle, C. M., & Hillman, C. H. (2003). Expert performance in sport: Current perspectives and critical issues. In J. L. Starkes & K. A. Ericsson (Eds.), *Expert performance in sports: Advances in research on sport expertise* (pp. 49–83). Champaign, IL: Human Kinetics.

Kitsantas, A., Zimmerman, B. J., & Cleary, T. (2000). The role of observation and emulation in the development of athletic self-regulation. *Journal of Educational Psychology*, *92*(4), 811–817. doi:10.1037/0022-0663.92.4.811

Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. doi:10.1177/1948550617697177

Monfardini, E., Gazzola, V., Boussaoud, D., Brovelli, A., Keysers, C., & Wicker, B. (2013). Vicarious neural processing of outcomes during observational learning. *PloS One.*, *8*(9), e73879. doi:10.1371/journal.pone.0073879

Moore, C. M., Lelievre, N., & Ste-Marie, D. M. (2019). Observing different model types interspersed with physical practice has no effect on consolidation or motor learning of an elbow flexion–extension task. *Human Movement Science*, *63*, 96–107. doi:10.1016/j.humov.2018.11.014

Pollock, B. J., & Lee, T. D. (1992). Effects of the model's skill level on observational motor learning. *Research Quarterly for Exercise and Sport*, *63*(1), 25–29. doi:10.1080/02701367.1992.10607553

Raedts, M., Van Steendam, E., De Grez, L., Hendrickx, J., & Masui, C. (2017). The effect of different types of video modelling on undergraduate students' motivation and learning in an academic writing course. *Journal of Writing Research*, *8*(3), 399–435. doi:10.17239/jowr-2017.08.03.01

Rak, N., Bellebaum, C., & Thoma, P. (2013). Empathy and feedback processing in active and observational learning. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(4), 869–884. doi:10.3758/s13415-013-0187-1

Ram, N., Riggs, S. M., Skaling, S., Landers, D. M., & McCullagh, P. (2007). A comparison of modelling and imagery in the acquisition and retention of motor skills. *Journal of Sports Sciences*, *25*(5), 587–597. doi:10.1080/02640410600947132

Rohbanfard, H., & Proteau, L. (2011). Learning through observation: A combination of expert and novice models favors learning. *Experimental Brain Research*, *215*(3–4), 183–197. doi:10.1007/s00221-011-2882-x

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. doi:10.1080/00222890009601357

Simonsohn, U. (2015). Small telescopes: Detectability and the evaluation of replication results. *Psychological Science*, *26*(5), 559–569. doi:10.1177/0956797614567341

Ste-Marie, D. M., Law, B., Rymal, A. M., Jenny, O., Hall, C., & McCullagh, P. (2012). Observation interventions for motor skill learning and performance: An applied model for the use of observation. *International Review of Sport and Exercise Psychology*, *5*(2), 145–176. doi:10.1080/1750984X.2012.665076

Weeks, D. L., & Anderson, L. P. (2000). The interaction of observational learning with overt practice: Effects on motor skill learning. *Acta Psychologica*, *104*(2), 259–271. doi:10.1016/S0001-6918(00)00039-1

Weir, P. L., & Leavitt, J. L. (1990). Effects of model's skill level and model's knowledge of results on the performance of a Dart throwing task. *Human Movement Science*, *9*(3–5), 369–383. doi:10.1016/0167-9457(90)90009-3

Wesch, N. N., Law, B., & Hall, C. R. (2007). The use of observational learning by athletes. *Journal of Sport Behavior*, *30*(2), 219.

Wulf, G., Shea, C., & Lewthwaite, R. (2010). Motor skill learning and performance: A review of influential factors. *Medical Education*, *44*(1), 75–84. doi:10.1111/j.1365-2923.2009.03421.x

Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology*, *94*(4), 660–668. doi:10.1037/0022-0663.94.4.660

Received May 23, 2019
Revised February 12, 2020
Accepted February 13, 2020