Human Movement Science 47 (2016) 231–239

Contents lists available at ScienceDirect

Human Movement Science

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

On the optimal degree of fluctuations in practice for motor learning

Ernst-Joachim Hossner *, Boris Käch, Jonas Enz

University of Bern, Switzerland

ARTICLE INFO

Article history:
Received 15 August 2014
Revised 9 June 2015
Accepted 12 June 2015
Available online 26 June 2015

PsychINFO classification:
2330
2343
3700

Keywords:
Structural learning
Two-state model
Differencial learning
Contextual interference
Augmented feedback

ABSTRACT

In human movement science, it is widely accepted that random practice generally enhances complex motor-skill learning compared to repetitive practice. In two experiments, a particular variability-related concept is put to empirical test, namely the concept of *differencial learning* (DL), which assumes (i) that learners should not be distracted from task-space exploration by corrections, and (ii) that learning is facilitated by large inter-trial fluctuations. In both experiments, the advantage of DL over repetitive learning was not statistically significant. Moreover, learning was more pronounced when participants either received corrections in addition to DL (Exp. 1) or practiced in an order in which differences between consecutive trials were relatively small (Exp. 2). These findings suggest that the positive DL effects reported in literature cannot be attributed to the reduction of feedback or to the increase of inter-trial fluctuations. These results are discussed in the light of the *structural-learning* approach and the *two-state model* of motor learning in which structure-related learning effects are distinguished from the capability to adapt to current changes.

1. Introduction

After the empirical evidence obtained in the wake of Schmidt's (1975) *schema theory* and *variability-of-practice hypothesis*, it seems to be widely accepted in human movement science that variable practice generally enhances complex motor-skill learning compared to constant, repetitive practice (for an early critical review, see van Rossum, 1990). In retrospect, however, it was only some years later that the schema-theoretical implication of restricting variability to the limits defined by the features of "generalized motor programs" could be overcome. In this regard, Shea and Morgan (1979) were able to show that practicing arm-movement patterns with different sequencing in a random order outperforms a training protocol where the same patterns are exercised in a blocked fashion. As both groups received exactly the same amount of practice and differed only in the sequential arrangement of the trials, it could be further inferred from this study that the advantage of variable practice cannot be solely attributed to rule abstraction from the overall gathered motor experience. Instead, the structure of the practice session and, in particular, the inter-trial variance also seemed to matter. Referring to the *contextual-interference* (CI) hypothesis, which was originally formulated by Battig (1972) for verbal learning, Shea and Morgan (1979, p. 179) related the revealed superiority of random practice to the idea that "contextual interference is closely associated with ... changes across trials in the experimental and processing contexts" and that "practice under increased

* Corresponding author at: University of Bern, Movement Science, Bremgartenstrasse 145, CH-3012 Bern, Switzerland.
E-mail address: ernst.hossner@ispw.unibe.ch (E.-J. Hossner).

http://dx.doi.org/10.1016/j.humov.2015.06.007
0167-9457/© 2015 The Authors. Published by Elsevier B.V.

contextual interference can produce more elaborate and distinctive processing of the material to be learned and thus facilitate delayed retention" (for a critical review of the CI effect in motor learning, see Brady, 1998).

More recently, an alternative explanation for the advantage of random over constant or blocked practice has been proposed in terms of *differencial learning* (DL). This approach was introduced by Schöllhorn (1999), who labelled this form of learning "differencial" instead of "differential" in full awareness of the spelling 'error' in order to stress the approach-specific importance of differen-c-es: "The label 'differencial' particularly emphasises the differences which are produced by two consecutive movements" (Schöllhorn, Beckmann, Janssen, & Michelbrink, 2009, p. 38). As it is assumed that "in all established training approaches, learning seems to be induced by the differences between two consecutive trials and less by identical movement parts" (Hegen & Schöllhorn, 2012, p. 32), it is expected "that an additional intensification of the fluctuations ... enhances performance" (Schöllhorn, 2005, p. 129). "Through a permanent process of creating differences between two consecutive movement executions, the always present noise in the movement is amplified in order to explore the potential task space and to automatically interpolate within these boundaries" (Schöllhorn, Sechelmann, Trockel, & Westers, 2004, p. 13). "Hence, the point is to increase the proportion which deviates from the preceding exercise in order to enhance learning" (Hegen & Schöllhorn, 2012, p. 33).[1]

On a theoretical level, Schöllhorn et al. (2006, p. 16) link the DL concept to the *stochastic-resonance* phenomenon (Ward, Neiman, & Moss, 2002), since "stochastic resonance might benefit training when two noisy signals interact in resonance, with one noisy signal equating to the continuously changing exercises or instructions and the other noisy signal formed by the noisy, dynamic condition of each athlete (e.g. during movement repetitions). ... [Hence,] by confronting [for instance] a footballer with a higher number of practice activities, the probability increases that any of the training exercises can get in resonance with the athlete's needs."

Empirically, the DL approach has been tested in experiments in which "a 'traditional' group practiced the skills under task constraints that contained little inter-trial variability, while another 'differential training' group practiced the skills with variability added to the target skill in the form of random, additional and irrelevant movement components" (Schöllhorn et al., 2006, p. 6). On the basis of such an experimental design, Schöllhorn et al. (2004), among others, obtained superior goal-shot precision, while Beckmann and Schöllhorn (2006) found superior shot put performance in post- and retention tests as a consequence of a differencial-learning intervention (see also Schöllhorn et al., 2006, or Savelsbergh, Kamper, Rabius, de Koning, & Schöllhorn, 2010).

Hence, one ends up with two alternative explanations of the same phenomenon: the CI approach focusing on the elaborateness of the abstracted representation and the DL approach focusing on the exploitation of stochastic resonance. At this point, regarding the design of a decisive empirical test, it has to be recognized that an advantage of variable over constant, random over blocked as well as differencial over traditional practice would be compatible with both accounts. Consequently, the problem arises that an increase of contextual interference is necessarily accompanied by an increase of exploitable inter-trial fluctuations and vice versa. Hence, it seems hard to empirically test both explanatory concepts against each other.

In trying to distinguish the two approaches in the context of motor learning, Schöllhorn et al. (2006) pointed at the fact that "alternating practice of two or more movements at the same time [in CI learning] only covers a small plane in a high dimensional task solution space, while the addition of randomized variability [in DL] covers a much wider area of the high dimensional task solution space" (p. 18). Thus, it can be inferred that, from a DL standpoint, a reduction of the degree of inter-trial variance that was assumed as optimal in classical DL experiments should deteriorate motor learning. As illustrated in Fig. 1, the empirical effects of such a reduction could be investigated by comparing a differencial-learning (DL) group with a structural-learning (SL) group on the one hand and, for control purposes, with a traditional-learning (TL) group on the other hand. Whereas DL participants would experience the whole task space with substantial fluctuations between trials, SL participants would practice exactly the same variants but with a minimization of inter-trial variance. As it can be assumed that, due to the structured order of the variants, the SL protocol still – or even better – allows for the abstraction of an elaborate representation of the task space, no learning disadvantage for the SL compared with the DL group should be predicted from a CI perspective. Hence, one ends up with an experimental between-group design that can claim to count as a decisive experiment. In the present study, this design was pursued in Experiment 2.

Before putting this comparison to empirical test, however, it seems useful to conduct a preliminary experiment to assess the role of augmented feedback in the context of DL. As sketched above, the DL protocol emanates from the idea that the participants should find their individual task-space optima by themselves as a result of the experimentally induced fluctuations. For this reason, participants are typically not given corrective feedback based on some allegedly ideal or optimal technique. In contrast, augmented feedback, which forms a crucial part of most standard practice programs, is typically given in the TL group. Consequently, in reference to the *guidance hypothesis* proposed by Schmidt (1991) and the benefits of reduced feedback frequency found by Winstein and Schmidt (1990), the reported superiority of the DL group might also be explained

[1] In this paragraph, the following citations have been translated from German to English by the authors: "Der Begriff des 'differenziellen' betont dabei insbesondere die Differenzen, die durch zwei aufeinander folgende Bewegungen erzeugt werden" (Schöllhorn et al., 2009, p. 38). "Danach scheint bei allen bekannten Trainingsansätzen das Lernen anhand der Differenzen zweier aufeinander folgender Bewegungen stattzufinden und weniger anhand der identischen Bewegungsanteile" (Hegen & Schöllhorn, 2012, p. 32). "...dass eine zusätzliche Verstärkung der ... Fluktuationen eine leistungssteigernde Wirkung besitzt" (Schöllhorn, 2005, p. 129). "Durch das ständige Erzeugen von Differenzen zwischen zwei aufeinander folgenden Bewegungsausführungen wird das stets vorhandene Rauschen in der Bewegung verstärkt, um den möglichen Lösungsraum abzutasten und innerhalb dieser Grenzen automatisch zu interpolieren" (Schöllhorn et al., 2004, p. 13). "Es geht somit darum, den von der vorigen Übung abweichenden Anteil zu vergrößern und so den Lernerfolg zu steigern" (Hegen & Schöllhorn, 2012, p. 33).

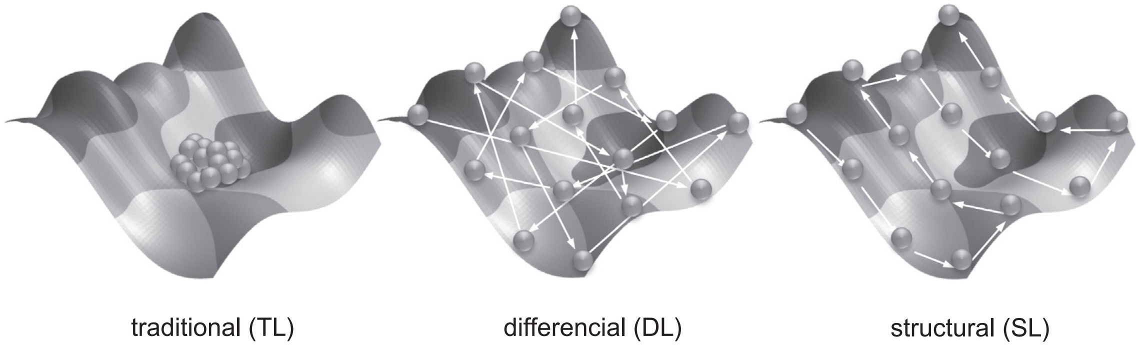

Fig. 1. Task-space exploration in a traditional-learning (TL, left), a differencial-learning (DL, middle) and a structural-learning (SL, right) intervention. DL and SL do not differ regarding the practised movement variants but solely regarding inter-trial variability.

by the reduction of augmented feedback. Although the omission of corrective feedback seems to be a by-product rather than an essential feature of the DL concept, it still seems worthwhile to disentangle this issue to make sure that effects found in the above sketched three-group design can solely be ascribed to differences in inter-trial variability and cannot alternatively be explained by differences in the amount of feedback received by the participants. Therefore, in Experiment 1, the DL and TL group was complemented by a DL + FB group whose participants practiced the same variants in the same order as the DL group but, unlike the latter group, were frequently provided with prescriptive technique-related augmented feedback.

2. Experiment 1

Experiment 1 resembles the Schöllhorn et al. (2004) study on goal-shot precision training in football. However, their design with a TL and DL group was extended by the addition of a DL + FB group, whose participants frequently received corrective feedback. From a DL – as well as from a CI – perspective, superior learning should be predicted for the DL compared to the TL group. Furthermore, if the absence of corrective feedback would be responsible for this superiority, worse post-test performance should be expected for the DL + FB participants compared to the DL participants. If, however, the DL + FB group would achieve similar or even better post-test results than the DL group, the superiority of the DL protocol to the TL group should be solely ascribed to structure-related specifics of the practice sessions.

2.1. Methods

2.1.1. Participants

Thirty male youth football players volunteered in the study. They were recruited from a D-junior team ($N = 14$, age: 12–13 years) and a C-junior team ($N = 16$, age: 14–15 years) of the Swiss football club FC Köniz. Due to personal issues, two participants dropped out in the course of the experiment, so that, in the end, a complete data set of 28 players was available. The participants were treated in accordance with the Declaration of Helsinki, implying that permission to participate was obtained from the parents in advance.

Participants were allocated to one of the three experimental groups on the basis of their pre-test results, also taking age ($M = 13.8$ years, $SD = 1.1$ years) and football experience (club membership, $M = 5.8$ years, $SD = 1.9$ years) into account. After two dropouts, data were available from 10 DL, 9 TL, and 9 DL + FB participants. In each group, D- and C-juniors were represented by a minimum of 3 and a maximum of 6 players. Some participants were not able to attend all 12 training sessions (times absent, $M = 0.9$, $SD = 1.1$). However, the groups did not differ significantly from each other in any of these variables (all $ps > .65$).

2.1.2. Procedure

The study was carried out on a football-practice venue where all participants were tested and treated in sub-groups at the same time. The intervention phase took 6 weeks with 2 sessions a week, each lasting 30 min. The players received a standardized warming-up of 20 min at the beginning and finished the session with a further 40 min of football training with their regular teams. Over the intervention phase, no specific goal-shot exercises were included in the regular team training. In each session, 30–35 goal shots were conducted per participant, resulting in an overall number of about 400 practice trials over the 12 treatment sessions. Three experienced football coaches who were not aware of the research question guided the sessions. After two weeks, the instructors rotated to another group, such that each participant practiced four times with each instructor and, due to the parallel conduction of the sessions, under exactly the same weather conditions.

The TL group received exercises that were structured in a methodologically ordered series with easy tasks at the beginning and an increase of task difficulty over the course of the sessions. Based on the recommendations of Mayer (2001) and Peters (2001), the exercises aimed at optimizing the sweet-spot kick first and, subsequently, at learning the inside and outside kick to be able to produce curved ball flights. In each session, 3–4 exercises were conducted, which were repeated 8–10 times each. Whenever procurable, participants received augmented feedback by the instructor, which was typically focused on the recent performance in relation to the ideal kicking technique. The players used official age-group specific balls, i.e. size 4 (circumference: 63.5–66.0 cm) balls for the D-juniors and size 5 (circumference: 68.5–70.0 cm) balls for the C-juniors.

As in the TL group, the DL intervention matched the respective treatment of the Schöllhorn et al. (2004) study as closely as possible. The DL variants can be assigned to 13 sources of variation that are characterized in more detail in Table 1. Over the six weeks of treatment, the training exercises focused on the supporting leg first (A and B), followed by variants concerning the kicking foot (C and D) and the trunk position (E and G). Finally, movement-related variants were complemented by variations regarding the approach (I and J), the kicking movement (K and L), and ball characteristics (M). Participants were instructed only regarding the specific variant. This means that, for instance, for the J1 variant, they were just advised to approach "on forefeet". However, in reference to the DL assumption that learners should find their individual movement optimum by themselves, participants were free to choose the speed and further specifics of the approach in accordance to personal preferences.

Initially, the variants were introduced in isolation and, in the next step, combined with other sources of variation until, at the end, three sources were demanded in combination. Furthermore, participants were never provided with additional augmented feedback regarding errors or their performance in relation to an optimal movement execution. Instead, the instructors gave feedback with respect to the degree the player successfully fulfilled the task at hand. In cases of an incorrect execution, the task was repeated up to a maximum of three times until the actual performance matched the demands. With the exception of these cases, over the whole treatment phase, participants were never assigned to the same task twice.

The only difference between the DL and the DL + FB groups referred to movement-related augmented feedback that was given in the DL + FB group only. In the DL + FB group, the optional task-related repetitions of the original DL intervention were particularly used to provide the participants with technique-related information. Furthermore, the instructor gave corrections individually whenever a non-optimal performance was noticed that could not be attributed to the current task variant. In addition, augmented feedback was given to the whole group after each task block. Consequently, each participant of the DL + FB group received technique-related corrections after about every third trial.

2.1.3. Measures

A pre-test was conducted one week before the first training session and a post-test one week after the last training session. Due to the retention interval of one week after the end of the acquisition phase, the post-test could also be regarded as a retention test, that is, a test that measures learning effects in terms of relatively permanent behavioral changes. In each test, after a collective warming-up of 20 min, the participants were tested one at a time. Each player had to perform 16 goal shots: 8 shots each from a left and right dead-ball position subdivided into 4 shots each onto a target in the left and right corner in the goal. For C-juniors, the balls were positioned either at the left or right interception point of the penalty-box line with the penalty arc, and, due to their younger age, from these points, 2 m closer to the goal line for D-juniors, resulting in a perpendicular distance to the goal of 16.5 m and 14.5 m, respectively. Beyond this facilitation, D-juniors used size 4 balls (circumference: 63.5–66.0 cm), whereas C-juniors had to shoot size 5 balls (circumference: 68.5–70.0 cm). Two red plastic discs (diameter: 0.2 m) served as targets that were each fixed with plastic cords in the upper left and right corner of the goal, with a distance of 0.5 m to the cross bar and 1.0 m to the goal post.

To enhance the reliability of the measurement, all test trials were filmed with two video cameras (Sony Handycam HDRXR520V, frame rate: 60 Hz) from a frontal and a lateral view, respectively. This approach allowed for a delayed computerized analysis of the synchronized tapes (Dartfish 4.5). For this purpose, the frame of the lateral-view tape was determined when the ball passed the goal plane in order to analyze the corresponding frame of the frontal-view tape regarding the current distance from the ball to the center of the target (radial error, in cm). The average of these measurements over the 16 trials served as the dependent variable of the experiment.

Shooting accuracy in pre- and posttest was tested on differences by an initial 3 (groups) $\times$ 2 (pre-/posttest) ANOVA with repeated measures on the second factor and, in the case of significant effects, by respective post hoc tests. The significance level for tests on differences was a priori fixed as $\alpha = .05$. The a priori definition of appropriate sample sizes was based on effect sizes reported in literature for similar group comparisons (Schöllhorn et al., 2004). A posteriori effect sizes are reported in the results section as η_p^2-values.

2.2. Results

In Fig. 2, the pre- and post-test results of the three groups are depicted. All groups improved shot accuracy from pre- to post-test, as indicated by a significant repeated-measures main effect of the 3 $\times$ 2-ANOVA, $F(1,25) = 38.60$, $p < .01$, $\eta_p^2 = .61$. The performance of the DL participants improved more than the performance of the TL participants, however, the relevant 2 $\times$ 2-ANOVA interaction was far from significant, $F(1,17) = 0.56$, $p = .46$, $\eta_p^2 = .03$. The same held true for the interactions of the other two 2 $\times$ 2-ANOVAs, that is, for the comparison of the TL and DL + FB group, $F(1,17) = 1.00$, $p = .33$, $\eta_p^2 = .06$, as well as for the comparison of the DL and DL + FB group, $F(1,17) = 0.04$, $p = .84$, $\eta_p^2 < .01$.

Table 1
Sources of variation for the goal-shot technique and task variants applied in the DL and DL + FB group of Experiment 1.

Source of variation		Variant 1	Variant 2	Variant 3
A	Supporting leg (I)	Behind the ball	*Regular*	In front of the ball
B	Supporting leg (II)	With knee bent	*Regular*	With knee extended
C	Kicking foot (I)	Bent to the ground	*Regular*	Parallel to the ground
D	Kicking foot (II)	Pronated	*Regular*	Extended
E	Trunk position (I)	Right-backward	Backward	Left-backward
F	Trunk position (II)	Right-forward	Forward	Left-forward
G	Trunk position (III)	To the right	*Regular*	To the left
H	Approach (I)	Accelerating	*Regular*	Decelerating
I	Approach (II)	With large steps	*Regular*	With small steps
J	Approach (III)	On forefeet	*Regular*	On heels
K	Kicking movement (I)	To the right	*Regular*	To the left
L	Kicking movement (II)	Stopped at contact	*Regular*	Full swing
M	Ball	Larger	*Regular*	Smaller

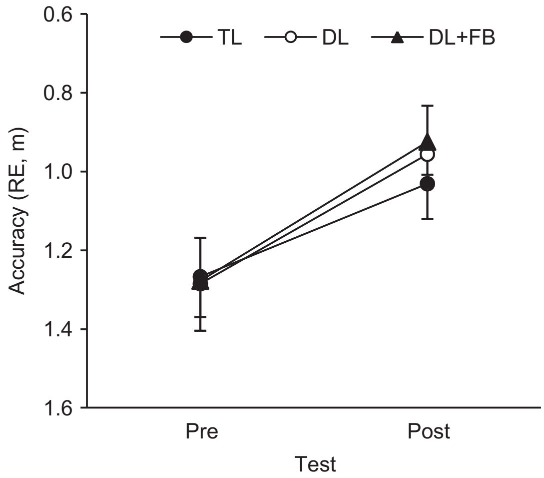

Fig. 2. Pre- and posttest performance (goal-shot accuracy) of the traditional-learning (TL), differencial-learning (DL) and differencial-learning-plus-feedback (DL + FB) groups in Experiment 1. Error bars represent *SE* and are displayed one-sided only.

2.3. Discussion

As indicated by the non-significant interaction for the comparison of the DL with the TL group, the significant advantage of DL reported by Schöllhorn et al. (2004) could not be replicated in the present experiment. However, the most interesting point of the study at hand refers to the fact that the addition of augmented feedback in the DL + FB group did not lead to a significantly worse learning effect than that of the DL group; in fact the improvement in performance of the DL + FB group was (albeit not significantly) larger than the improvements of the two other groups. From this finding it can be at least inferred that augmented feedback did not hinder learning. Hence, it seems that the advantage of DL reported in the literature must have been due to the structure of the practice sessions, in particular, to the inter-trial variability induced by the DL protocol. Therefore, we are confident that possible effects of the latter variable in Experiment 2 will not be confounded by differential effects of the feedback provided to the three experimental groups.

3. Experiment 2

Experiment 2 was designed in reference to the shot put experiment conducted by Beckmann and Schöllhorn (2006). However, following the argument made in the introduction (see Fig. 1), a third, SL, group was added to the original TL and DL group. This SL group practiced the same variants as the DL group, but in a structured order, that is, with a minimization of differences between consecutive trials. From a DL perspective, superior learning should be expected for DL in comparison with both, SL and TL. In contrast, the empirical finding of a comparable or even superior retention-test performance of the SL in comparison with the DL group would be taken as contradictory evidence for the DL approach.

3.1. Methods

3.1.1. Participants

Participants were treated in accordance with the Declaration of Helsinki. In total, 39 right-handed sport-science students of the University of Bern (15 female, 24 male) were recruited for participation. They received course credit in return. Due to either illness or injuries, three students had to withdraw from the study, such that a complete data set could be obtained for 36 participants (13 female, 23 male), that is, 12 participants per group.

Participants were allocated to one of three experimental groups on the basis of their pretest results, also taking the following variables into account: age (M = 20.9 years, SD = 0.9 years), sex (initially 5, finally 4 or 5 females per group), body height (M = 175.2 cm, SD = 8.1 cm), shot put experience (4-point Likert scale with 4 = extensive, M = 1.2, SD = 0.4), and their motivation to take part in study (4-point Likert scale with 4 = very high, M = 2.3, SD = 1.1). Due to other obligations, some participants could not attend all 8 training sessions (times absent, M = 0.7, SD = 0.7). The three groups did not differ significantly from each other on any of these variables (all $ps > .40$).

3.1.2. Procedure

The whole study was conducted indoors in a University sports hall. The participants received 8 sessions of a group-specific shot put training of 50 min each, distributed over 4 weeks. After a couple of warm-up exercises, Sessions 1–7 consisted of 32 practice trials each and Session 8 of 20 practice trials that were immediately followed by the post-test. Hence, the total number of practice trials was 244. Training sessions were conducted on days 5 and 6 of each intervention week, beginning at 7.50, 8.30, and 9.10 am, with warm-up phases of 10 min in a separate part of the gym. Due to a daily rotation of the schedule, on average, all participants conducted their training program at the same time of day. An experienced track-and-field coach, who was not aware of the research question, led all training sessions.

The treatment of the TL group was structured in accordance with the advice given by Bauersfeld and Schröter (1998) and Jonath, Krempel, Haag, and Müller (1995), two classical textbooks in German sports practice literature on teaching and learning in track and field athletics. The teaching approach was characterized by an initial decomposition of the target movement and a methodologically ordered series of exercises that were related to the crucial elements first and to their successive combination afterwards until, finally, the standing technique of the shot put was practiced as a whole. Depending on the position in the series, each exercise was repeated 4, 6, 8, or 12 times, and the experimenter accompanied the learning process by giving frequent instructions and corrections. Due to a limited availability of indoor shots of the same weight, 4.00 kg shots were used for the female and 6.25 kg shots for the male participants.

As depicted in Table 2, for the DL and SL interventions, seven sources of movement variation were defined with either two (A, B) or four (C, D, E, F, G) variants each, which had been derived from the regular target movement. In variant B2, a lighter shot weight had to be specified with respect to sex so that female participants practiced with 3.00 kg shots and male participants with 4.00 kg shots. In each practice trial, two sources of variation were combined to define a certain movement task variant. Thus, for instance, a certain task variant could be to perform the shot put with the left arm and to conduct the leg movement more slowly than normal (A2–G1), or to start the movement with the shot close to the opposite shoulder and to put it to the left with a putting angle larger than regular (D4–F2). This approach resulted in a total of 244 task variants that were presented exactly once over the whole intervention. At the beginning of each training session, the sources of variation and the respective variants were explained to the participants using illustrations. On this basis, it sufficed when, before each trial, the next task variant was assigned to the participants just by presenting the two pictures that separately illustrated the combination of two movement variants. No further instructions and, in particular, no movement-related corrections were given to the participants.

The only difference between the DL and the SL groups referred to the order in which the 244 task variants were presented to the participants. For the DL group, the 244 variants were randomized, resulting in comparatively large differences from one trial to the next. In contrast, these differences were minimized in the SL group as each source's variants were presented in succession without changing the variant in combination. An order obeying this rule would be A1–B1, A1–B2, A1–B3, A1–B4, A2–B4, A2–B3, A2–B2, A2–B1, A3–B1, etc. In this order, the A-variant is kept constant for the first four trials (A1) while the four B-variants are presented in a structured order (from B1 to B4). Subsequently, from trial 4 to trial 5, only the A-variant changes (from A1 to A2) whereas the B-variant B4 is kept constant and from trial 5 to trial 8, the B-variants are presented in a structured order again (from B4 to B1) with an unchanged A-variant (A2). Consequently, the participants of the SL group experienced differences from trial to trial; however, the inter-trial variance was minimized.

3.1.3. Measures

Participants were tested one week before the first training session in a pre-test, at the end of the last session in a post-test, and two and four weeks after the last session in retention tests 1 and 2, respectively. Female participants executed the shot puts with a 4.00 kg shot and male participants with a 6.25 kg shot. The pre- and retention tests were conducted after warming-up periods of 10 min, which consisted, as usual in sports practice, of a general physical activation including stretching exercises and a more specific part in which three shot puts had to be performed. The participants were informed that these warm-up trials were not included in the test score. As the post-test was undertaken at the end of the last practice session, no further warm-up was carried out in this case. The tests were conducted in groups of 2–4, so that the participants were provided with sufficient recovery time between the test trials. Two experimenters measured performance after each

Table 2
Sources of variation for the shot put technique and task variants applied in the DL and SL group of Experiment 2.

Source of variation		Variant 1	Variant 2	Variant 3	Variant 4
A	Active arm	*Regular*	Left		
B	Shot weight	*Regular*	Lighter		
C	Rear foot	Slightly back	*Regular*	Slightly forward	Quite forward
D	Shot position	Shoulder	*Regular*	Sternum	Opposite shoulder
E	Feet position	Minimal	Hip-width	*Regular*	Maximal
F	Put direction	Upper-right	Upper-left	*Regular*	Lower-middle
G	Deceleration	Leg movement	Trunk movement	Arm movement	*Regular*

individual trial using a standard measuring tape (in cm). The average distance achieved in the three test trials served as dependent variable of the experiment.

Changes from pre- to post- and retention tests were approached by the logic of a 3 (groups) $\times$ 4 (pre-/post-/retention test 1 and 2) ANOVA with repeated measures on the second factor. The a priori definition of the overall sample size and the significance level for tests on differences ($\alpha = .05$) was based on effect sizes reported in literature for similar group comparisons (Beckmann & Schöllhorn, 2006). In the following, a posteriori effect sizes will be reported as η_p^2-values.

3.2. Results

The results of the three groups in pre-, post- and retention tests are depicted in Fig. 3. A one-way ANOVA on the pretest results showed that the three groups did not differ significantly on the pretest, $F(2,33) = 0.20$, $p = .82$, $\eta_p^2 = .01$. Moreover, given that the predictions refer to performance improvements small differences in pretest levels do not pose a problem.

From pre- to post- and retention tests, descriptively, the TL group stays on the same performance level, whereas the DL group improves slightly and the SL group considerably. Within the three groups, no remarkable development can be determined from post- to retention tests 1 and 2, underlined by the fact that a respective 3 $\times$ 3 ANOVA revealed neither a significant repeated-measures effect, $F(2,66) = 0.93$, $p = .40$, $\eta_p^2 = .03$, nor a significant interaction, $F(4,66) = 0.36$, $p = .84$, $\eta_p^2 = .02$. Thus, when it comes to inferential statistics, it seems most reasonable to test pre-test results against the mean of post- and retention test scores. In this regard, a 2 $\times$ 2 ANOVA on the difference between TL and DL revealed no significant interaction, $F(1,22) = 2.19$, $p = .15$, $\eta_p^2 = .09$. In contrast, a significant interaction was obtained for the TL–SL comparison, $F(1,22) = 4.29$, $p = .05$, $\eta_p^2 = .16$, whereas the TL–DL interaction failed to reach significance, $F(1,22) = 1.31$, $p = .27$, $\eta_p^2 = .06$.

3.3. Discussion

As the learning advantage of DL over TL could be confirmed by trend only, the replication of the Beckmann and Schöllhorn (2006) study failed in inferential-statistical respect. In contrast, the superiority of the SL intervention was underpinned by the only statistically significant effect present, namely, the significant TL–SL interaction. Thus, the pattern of results is at odds with the most fundamental assumption of the DL approach, that is, the assumption that large differences between consecutive trials improve learning. Instead, the results favor the assumption that the elaboration of task variants is optimized by the application of a training protocol that is variable on the one hand, but structured on the other hand.

4. Overall discussion

In sum, the results of Experiment 1 indicated that the advantage of random over traditional practice cannot be ascribed to the – in DL studies concomitant – reduction of corrective feedback, while the results of Experiment 2 suggested that the random-practice effect should be attributed to the elaboration of task variants rather than to the exploitation of inter-trial fluctuations.

To the best of our knowledge, Experiment 2 is the only empirical study to date that focuses on the comparison of structured vs. random practice in complex motor-skill learning. However, with respect to the line of argument pursued here, the results obtained in two other groups of experiments might be taken as converging evidence. The first group refers to the *errorless-learning hypothesis* proposed by Maxwell, Masters, Kerr, and Weedon (2001). In their golf putt learning study (Experiment 1), an order of variable-practice trials in which errors were minimized (distances to the hole monotonically increasing from 0.25 to 2.00 m) was compared not only with an order in which errors were maximized (distances monotonically decreasing from 2.00 to 0.25 m) but also with a random order (distances randomly varying within the range of 0.25–2.00 m). In the context at hand, the gradually changing experiences of the two error-related interventions can be understood as SL, whereas the increased inter-trial differences qualify the random protocol as DL (to a more pronounced degree at least). Due to their main research question, the authors only report on a significant advantage of the errorless-learning group in transfer and retention tests in comparison with the two other groups. In all tests, however, at least descriptively, not only the protocol for errorless learning but also the error-rich-learning intervention resulted in better learning than random

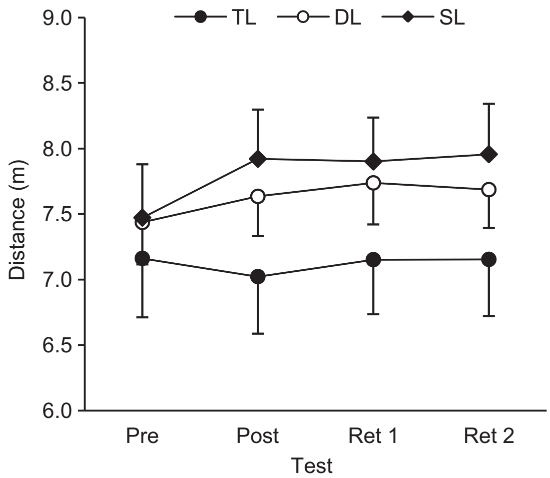

Fig. 3. Pre-, post- and retention test performance (shot put distance) of the traditional-learning (TL), differencial-learning (DL) and structured-learning (SL) groups in Experiment 2. Error bars represent *SE* and are displayed one-sided only.

practice (Maxwell et al., 2001, p. 1055, Fig. 2). Thus, transferred to the research question at hand, it can be said that SL led to an overall better learning than DL.

The second set of evidence stems from the domain of *structural learning* (Braun, Mehring, & Wolpert, 2010). In this respect, Turnham, Braun, and Wolpert (2012) exposed participants to 100 visuo-motor rotation conditions, which differed with regard to the angle of the input–output mapping. These variants were presented in blocks of 16 trials. However, while participants of the two gradual-learning groups experienced the variants in a structured order, the mapping angle changed quasi-randomly from block to block in the random-learning group. Overall, the "results suggest that rapid changes ... are not necessary conditions for facilitation of learning ... Indeed, rapid changes in task parameters might reduce later adaptation rates ... [On the contrary,] gradual subjects adapt more fully to the training rotations. This may lead to more thorough learning of the structure" (Turnham et al., 2012, p. 1119). Hence, with respect to permanent behavioral change, a SL intervention again turned out to be superior over a random protocol that can be interpreted as DL.

On the basis of these findings, it should be acknowledged that the advantage of structural over differencial practice found in Experiment 2 is corroborated by empirical results obtained in other fields of motor-learning research. Hence, when it comes to relatively permanent changes in motor behavior, it can be concluded that a more structured learning program seems to be superior over an intervention that is featured by pronounced inter-trial fluctuations.

However, motor learning does not only refer to the degree a movement pattern can be optimally reproduced; another characteristic of motor expertise concerns the capability to quickly adapt to current changes. In this regard, Smith, Ghazizadeh, and Shadmer (2006) have proposed a *two-state model* for motor learning. According to this model, "learning appears to be supported by a fast adaptive process that is highly sensitive to error but has poor retention and a slow adaptive process that has poor sensitivity to error but good retention" (Huang & Shadmer, 2009, p. 931). At this point, it is important to amend that, in their above-sketched force field learning study, Turnham et al. (2012) not only found advantages of gradual over random practice with respect to a better learning of the structure; in addition, "subjects that experienced random training showed a significantly lower movement initiation time compared with subjects who experienced gradual or no training" (Turnham et al., 2012, p. 1121). These lower initiation times of the participants with random practice might be interpreted as a superior capability to quickly adapt to current changes. Thus, it seems likely to relate SL to slow and DL to fast processes as proposed by Smith et al. (2006).

If this consideration is well grounded, the efficacy of DL should actually *not* be tested by performance parameters such as goal shot precision (Schöllhorn et al., 2004) or shot put distance (Beckmann & Schöllhorn, 2006). Instead, it seems advisable to assess DL effects by the degree of retained motor performance in situations that require a fast adaptation to situational changes as it would be the case, for instance, in windsurfing under permanently changing wind conditions. For the retention of complex motor skills per se, however, the application of a more structured learning protocol should be preferred.

Acknowledgments

The authors like to thank Nicole Naef, Gabriel Sager and Alexander Steiger for their assistance in the goal-shot study and Reto Büchi for his support as an instructor and as a designer of the traditional-learning intervention of the shot put experiment.

References

Battig, W. F. (1972). Intratask interference as a source of facilitation in transfer and retention. In R. F. Thompson & J. F. Voss (Eds.), *Topics in learning and performance* (pp. 131–159). New York: Academic Press.

Bauersfeld, K.-H., & Schröter, G. (1998). *Grundlagen der Leichtathletik [Foundations of track-and-field athletics]* (5th ed.). Berlin/Germany: Sportverlag.

Beckmann, H., & Schöllhorn, W. I. (2006). Differenzielles Lernen im Kugelstoßen [Differencial learning of the shot put]. *Leistungssport, 36*(4), 44–50.

Brady, F. (1998). A theoretical and empirical review of the contextual interference effect and the learning of motor skills. *Quest, 50*, 266–293.

Braun, D. A., Mehring, C., & Wolpert, D. M. (2010). Structure learning in action. *Behavioural Brain Research, 206*, 157–165.

Hegen, P., & Schöllhorn, W. I. (2012). Gleichzeitig in verschiedenen Bereichen besser werden, ohne zu wiederholen? Paralleles differenzielles Training von zwei Techniken im Fußball [Improving simultaneously in two domains without repetitions? Parallel differencial training of two techniques in football]. *Leistungssport, 42*(3), 17–23.

Huang, V. S., & Shadmer, R. (2009). Persistance of motor memories reflects statistics of the learning event. *Journal of Neurophysiology, 102*, 931–940.

Jonath, U., Krempel, R., Haag, E., & Müller, H. (1995). *Leichtathletik 3. Werfen und Mehrkampf [Track-and-field athletics 3: Throwing and combined events].* Reinbek/Germany: Rowohlt.

Maxwell, J. P., Masters, R. S. W., Kerr, E., & Weedon, E. (2001). The implicit benefit of learning without errors. *Quarterly Journal of Experimental Psychology, 54A*, 1049–1068.

Mayer, R. (2001). *Torschusstraining Fußball [Goal-shot training football].* Reinbek/Germany: Rowohlt.

Peters, R. (2001). Tipps zum Trainerverhalten am Beispiel des Schusstrainings [Tips for coaches' behaviour using the example of shot training]. *Fußballtraining, 19*(7), 30–38.

Savelsbergh, G. J. P., Kamper, W., Rabius, J., de Koning, J., & Schöllhorn, W. (2010). New methods to learn to start in speed skating. A differencial learning approach. *International Journal of Sport Psychology, 41*, 415–427.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*, 225–260.

Schmidt, R. A. (1991). Frequent augmented feedback can degrade learning: Evidence and interpretations. In G. E. Stelmach & J. Requin (Eds.), *Tutorials in motor neuroscience* (pp. 59–75). Dordrecht/Netherlands: Kluwer.

Schöllhorn, W. I. (1999). Individualität – ein vernachlässigter Parameter [Individuality – a neglected parameter]? *Leistungssport, 29*(2), 5–12.

Schöllhorn, W. I., Beckmann, H., Janssen, D., & Michelbrink, M. (2009). Differenzielles Lehren und Lernen im Sport. Ein alternativer Ansatz für einen effektiven Schulsportunterricht [Differencial teaching and learning in sports. An alternative approach for effective teaching in physical education]. *Sportunterricht, 58*(2), 36–40.

Schöllhorn, W. I. (2005). Differenzielles Lehren und Lernen von Bewegung Durch veränderte Annahmen zu neuen Konsequenzen [Differencial teaching and learning of movement. Through changed assumptions to novel consequences]. In H. Gabler, U. Göhner, & F. Schiebl (Eds.), *Zur Vernetzung von Forschung und Lehre in Biomechanik, Sportmotorik und Trainingswissenschaft [On cross-linking research and teaching in biomechanics, motor-control and exercise science]* (pp. 125–135). Hamburg/Germany: Czwalina.

Schöllhorn, W. I., Michelbrink, M., Beckmann, H., Trockel, M., Sechelmann, M., & Davids, K. (2006). Does noise provide a basis for the unification of motor learning theories? *International Journal of Sport Psychology, 37*, 34–42.

Schöllhorn, W. I., Sechelmann, M., Trockel, M., & Westers, R. (2004). Nie das Richtige trainieren, um richtig zu spielen [Never practising the proper thing in order to play properly]. *Leistungssport, 34*(5), 13–17.

Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Learning, Memory and Cognition, 5*, 179–187.

Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biology, 4*, e179.

Turnham, E. J. A., Braun, D. A., & Wolpert, D. M. (2012). Facilitation of learning induced by both random and gradual visuomotor task variation. *Journal of Neurophysiology, 107*, 1111–1122.

van Rossum, J. (1990). Schmidt's schema theory. The empirical base of the variability of practice hypothesis: A critical analysis. *Human Movement Science, 9*, 387–435.

Ward, L. M., Neiman, A., & Moss, F. (2002). Stochastic resonance in psychophysics and in animal behavior. *Biological Cybernetics, 87*, 91–101.

Winstein, C. J., & Schmidt, R. A. (1990). Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 677–691.