



Dissociative effects of normative feedback on motor automaticity and motor accuracy in learning an arm movement sequence

Christina Zobe^a, Daniel Krause^{a,*}, Klaus Blischke^b

^a Department of Sports and Health, Paderborn University, Germany

^b Institute of Sport Science, Saarland University, Germany

ARTICLE INFO

Keywords:

Augmented feedback
Automaticity
Dual task
Motor learning

ABSTRACT

Within a pre-post-design, we scrutinized the effects of normative augmented feedback with positive and negative valence on learning motor accuracy, consistency as well as automaticity by means of a dual-task paradigm. Forty-two healthy physical education students were instructed to produce an arm-movement sequence as precisely as possible with regard to three spatial reversal points within a time limit of 1200 ms. Twenty-eight practiced an elbow-extension-flexion-sequence (690 trials) and 14 participants were tested as a control group without feedback practice. Valence of normative feedback was systematically manipulated by means of reference lines in a visual feedback display. The reference lines indicated performance of a putative peer-group either to be superior (negative valence, *Normative-Negative-Group*) or inferior (positive valence, *Normative-Positive-Group*) to participants' actual performance.

As a result, dual-task costs (*n-back error*) significantly decreased solely in the *Normative-Positive-Group*, $p = .003$, $\eta_p^2 = .51$, but in no other group. Surprisingly, the mean absolute error for the motor task significantly decreased (i.e., precision increased) only in the *Normative-Negative-Group* with a large effect size, but in none of the other groups. Motor consistency was not significantly affected by the valence of normative feedback. According to the hypotheses of error-provoked attentional control, positive feedback-valence appears to enhance skill automatization, while – unexpectedly – only negative feedback-valence seems to enhance movement precision, which may be explained by effects of feedback valence on the learners aspiration level.

1. Introduction

The result of an action is subjectively classified as failure or success, dependent on the individual's level of aspiration, the subjective expectation to his own future achievement in a given task (Heckhausen & Heckhausen, 2018, p. 196–198). In motor learning additional information on the movement itself or its terminal effect (i.e., augmented feedback) is provided to support learning. Feedback that deviates from the level of aspiration, signaling performance being worse (i.e., negative prediction error) or better (i.e., positive prediction error) than expected, leads to different modes of information processing and hence learning.

Feedback signaling performance that is better than expected is a rewarding stimulus for the learner. Based on the Reward-Prediction-Error-Hypothesis-of-Dopamine (Glimcher, 2011), Krause, Agethen, and Zobe (2018) proposed the Error-Provoked-Attentional-Control-Hypothesis (EPAC-Hypothesis) in the context of motor learning. According to this hypothesis, feedback with negative valence (i.e., error feedback like “your elbow was not extended enough”) provokes an increase of attentional control, which is

* Corresponding author at: Department of Sports and Health, Paderborn University, Warburgerstr. 100, 33098 Paderborn, Germany.
E-mail address: daniel.krause@upb.de (D. Krause).

<https://doi.org/10.1016/j.humov.2019.06.004>

Received 2 November 2018; Received in revised form 6 June 2019; Accepted 6 June 2019

Available online 17 June 2019

0167-9457/ © 2019 Elsevier B.V. All rights reserved.

mediated by a decrease of the firing rate of dopaminergic mid-brain neurons (Holroyd & Coles, 2002), and thus impedes motor automatization. Instead feedback with positive valence (i.e., rewarding feedback like “your elbow was extended really good”) induces long-term potentiation of those still activated neuronal networks that effected the rewarding event mediated by an increase in firing rate of dopaminergic mid-brain neurons (Glimcher, 2011). This should effectively promote progress from early stages of learning with control of movement sequences being dominated by attention-dependent control processes (e.g., based on a spatial code) to later stages of learning where attention-independent representations (e.g., based on a motor code) are established (Hikosaka et al., 1999, Hikosaka, Nakamura, Sakai, & Nakahara, 2002). This later stage is also characterized as the automatic execution stage (Chein & Schneider, 2012). These considerations are also supported by several neurophysiological findings showing involvement of attention-related prefrontal brain areas (e.g., anterior cingulate cortex) subsequent to feedback with negative valence (feedback related negativity; Botvinick, 2007; Hikosaka & Isoda, 2010; Holroyd & Coles, 2002; Seidler, Kwak, Fling, & Bernard, 2013).

Consistent with the EPAC-Hypothesis, recent research revealed that lower error feedback frequencies allow motor automatization to occur to a greater extent than higher error feedback frequencies in learning arm movement sequences. According to Krause et al. (2018), an experimental group with error feedback on no more than 14% of a total of 760 practice trials showed substantial dual-task cost reductions as an indicator of motor automaticity, while an experimental group with 100% error feedback did not show any such dual-task cost reductions. In a comparable experimental setting, Agethen and Krause (2016) also demonstrated provision of feedback with positive valence to support motor automatization. Here, positive valence was induced by means of a bandwidth feedback condition. That is, error feedback is provided only if the performance error exceeds a certain range, whilst performance within this range signals proper achievement of the movement goal. An alternative way to manipulate feedback valence, termed social-comparative or normative feedback, is to relate the learner's performance to the performance of others (e.g., peer groups performance). In the respective studies, values of putative peer groups are experimentally manipulated in order to systematically induce either a positive or a negative valence of normative feedback.

Normative feedback (as quantitative feedback that is augmented with a normative reference) has been shown to affect motor learning of underhand bean bag throwing (Ávila, Chiviawsky, Wulf, & Lewthwaite, 2012), balancing on a stabilometer (Lewthwaite & Wulf, 2010), and performing a temporal pattern of a sequential key pressing task (Wulf, Chiviawsky, & Lewthwaite, 2010). In these experiments, learners showed better retention performance when provided with positive normative feedback, suggesting their performance being superior to the performance of others. Negative normative feedback, however, is thought to induce a tendency of being more self-focused, causing subjects perhaps to adopt a more attentional mode of control (Wulf & Lewthwaite, 2016) and thereby, according to our above considerations, impede motor automatization. However, substantial empirical evidence that negative normative feedback indeed induces a more attentional mode of control and at the same time prevents movement automatization is still missing. Higher frequencies of movement corrections (mean power frequency in angle-time curve) in a balance task on a stabilometer with positive normative feedback have been interpreted to result from a higher degree of motor automaticity (Lewthwaite & Wulf, 2010). Those were apparent during practice sessions with feedback manipulations and a retention session. Lewthwaite and Wulf (2010) did not integrate more directly related measures of motor automaticity to verify their assumption of a higher degree in automaticity. Ong and Hodges (2017) performed a replication study but failed to find the differential effects on balance performance and mean power frequency. In addition, they integrated a secondary task (i.e., tone counting) without finding any differences between conditions with negative and positive normative feedback. Both studies show limitations as there are no pre-tests integrated in the study design. Further research is needed to advance the knowledge under which conditions normative feedback with varied valence has an impact on motor learning and especially motor automatization.

According to the described theoretical assumptions and empirical findings, we set up the following hypotheses.

Hypothesis 1 (accuracy). Practice with positive normative feedback leads to stronger effects in learning movement accuracy compared to practice with negative normative feedback.

Hypothesis 2 (consistency). Practice with positive normative feedback leads to stronger effects in learning movement consistency compared to practice with negative normative feedback.

Hypothesis 3 (automatization). Practice with positive normative feedback leads to stronger effects in acquiring motor automaticity compared to practice with negative normative feedback.

To test these hypotheses, we conducted an experiment intended to systematically dissociating the learning effects of positive and negative normative feedback, while incorporating a dual-task paradigm to assess motor automatization. That is, reduction of dual-task costs in the course of practice was taken as an indicator for the attentional demands initially required for controlling execution of the primary motor task to decrease in the course of progressive movement automatization. All hypotheses will be tested in a pre-post design with delayed retention tests (at least one night of sleep) in order to test learning understood as relatively permanent changes in performance (Schmidt & Lee, 2011, p. 327).

In addition, we will analyze the practice data to describe the short-term effects of different feedback approaches during the acquisition phase in more detail.

2. Methods

2.1. Participants

A total of 42 participants (20 female and 22 male, mean age = 22.5 years, $SD = 2.8$; 36 right-handed, 2 ambidextrous and 4 left-

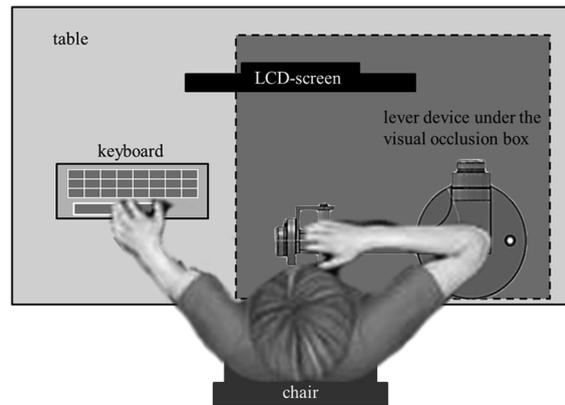


Fig. 1. Top view of a participant in the experimental set up with arm-lever device, visual occlusion, monitor and keyboard.

handed; German Version of the Edinburgh Handedness Inventory; Büsch, Hagemann, & Bender, 2010) participated in this study. The study has been conducted according to the Declaration of Helsinki and was approved by the ethics committee of the German Psychological Society (DGPs).

2.2. Apparatus and task

The apparatus was an underarm lever device, which allows movement in the horizontal plane. A visual mask occluded online movement control (Fig. 1). The lever arm was adjustable in length to adapt to individual lower arm lengths. A potentiometer was attached to the vertical axis functioning as a goniometer. The measurement chain consisted of a Novotech 6500 P6501 linear potentiometer (resolution: 0,007% Linearity: $\pm 0,05\%$, manipulating speed: 10 k rpm), an Advantech USB 4716 analog–digital-converter (16 bit 1000 HZ), a Voltkraft PS 1152A regulated power supply at 4.00 V with galvanic separation and a computer with DasyLab 10.0 plus Java-based software for data collection (movement time, movement reversals). Further experimental setup components were a monitor (17"; 4:3) installed above the lever device for the secondary task and feedback presentation, an adjustable chair for replicable positioning of the subjects in vertical, horizontal, and sagittal planes, and a keyboard on the left side of the lever as the input device for the secondary task response.

The criterion task was an elbow-extension-flexion sequence with three movement reversals at 70° , 20° , and 70° , measured from a defined starting position (0°) of the arm lever device. The participants were instructed to perform the task as precisely as possible and with a movement time below 1200 ms. No goal was specified for relative timing.

The secondary task for assessing the dual-task costs was a visual spatial n -back task. In the utilized n -back task, rectangular boxes were presented in a 3×3 matrix in pseudorandomized order. During one trial, 9 stimuli were presented with a stimulus onset interval of 700 ms or 900 ms. The subjects had to compare each stimulus with the stimulus presented n (2 in the utilized task) stimuli before the current presentation. Each trial contained 2–4 n -back events (*matches*). The subjects had to confirm a *match* (accordance between the actual and the n -back stimulus) by pressing the spacebar with their left hand on the computer keyboard.

2.3. Procedure

The experiment was divided into seven experimental sessions for the interventional groups. During the first session (session 0, acquisition), all participants were positioned in front of the lever with the right underarm on the lever in the 0° position and an upper-arm-frontal-plane-angle of 130° (Fig. 1).

Seating height, as well as horizontal, and sagittal seat position were measured to ensure a replicable position. Participants were instructed about the procedure and the tasks via presentation slides on the monitor and standardized verbal comments of the experimenter. Participants had five trials consisting of observing a video model, followed by a task execution with terminal feedback on movement time in milliseconds and the deviation from the movement reversals, as a bar diagram with one bar for each of the reversals indicating the direction and magnitude of the error.

After these initial trials, the participants completed the *Pre-Acquisition-Test*. The test procedure for all tests (*Pre-Acquisition*, *Pre-Test*, *Retention*) consisted of the movement task and the visual-spatial n -back task (inter stimulus interval: 700/900 ms pseudo-randomized; stimulus duration: 500 ms, Fig. 2), under single- and dual-task conditions (6 trials each).

To reduce sequence effects, the tests were ordered as follows: three single-task trials of the secondary task, three single-task trials of the motor task, six dual-task trials, additional three single-task trials of the motor task and additional three single-task trials of the secondary task. Priority was instructed for the motor task to ensure undisturbed motor task execution and the possibility to measure dual-task costs with reference to the secondary task performance. The *Pre-Acquisition-Test* was followed by 40 acquisition trials with feedback for movement reversals and movement time. The initial acquisition phase should help to ensure movement times below 1200 ms, and the development of an initial attention dependent abstract spatial representation of the motor task before the *Pre-Test*.

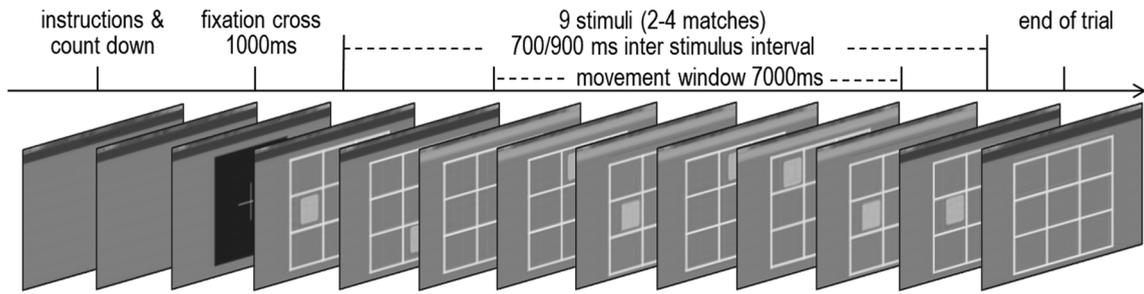


Fig. 2. Dual-task procedure with *n*-back stimuli.

The *Pre-Test* is relevant for testing the hypotheses, in that any possible treatment effects are reflected in the performance changes from *Pre-Test* to *Retention-Test*.

After the *Pre-Test* in session 1, subjects were assigned to two intervention groups and one passive control group parallelized according to pre-test results, gender and handedness: *Normative-Positive-Group*, *Normative-Negative-Group*, and *Passive-Control-Group* (each *n* = 14). During sessions 1–5 the two intervention groups underwent 690 trials of feedback practice (114 practice trials in session 1, 144 practice trials in session 2–5).

Augmented feedback regarding movement precision was provided on a computer screen on 14% of the feedback practice trials following a fading schedule (see Krause et al., 2018) in the two intervention groups. Errors were displayed as error bars on the monitor indicating reversal angles that were larger (bars above the zero line) or smaller (bars below the zero line) than the required angles of the target reversals. In both the *Normative-Positive-Group* and the *Normative-Negative-Group* white reference lines were overlaid (Fig. 3). These lines were systematically manipulated as to induce participants to assume their performance to be either superior (*Normative-Positive-Group*) or inferior (*Normative-Negative-Group*) to the putative peer-groups’ performance at the respective learning stage. Therefore, the participants were informed that the distance between those lines would change over the course of practice. Actually the distance of the reference lines for each subject was updated each six trials as that individual’s performance on the preceding block of 6 trials, multiplied by 1.5 (*Normative-Positive-Group*) or 0.4 (*Normative-Negative-Group*) was set as the new reference value. It is important to notice, that the error bars always provided the veridical quantitative error information in both groups whenever feedback was displayed. The white reference lines set an objective/nominal valence in addition to the quantitative error information. The resulting subjective/functional valence was measured by a questionnaire as it is described below.

The *Retention* was measured in session 6 with a retention interval of 48–72 h after the last practice session. The *Passive-Control-Group* did only accomplish the test-phases but not the practice sessions in session 1–5. This group was introduced to control for task integration as an alternative explanation for dual-task cost reduction (Blichke, Wagner, Zehren, & Brueckner, 2010; Ruthruff, Van Sleest, Johnston, & Remington, 2006).

Pre-Test and *Retention-Test* (factor *TIME*) required participants to perform the motor task (i.e. the primary task) and a visual spatial *n*-back task (i.e. the secondary task) under single-task (*ST*) and under dual-task (*DT*) conditions (factor *STDT*; six trials per condition) without feedback on either task performance. Participants were instructed to prioritize the primary task under dual-task conditions.

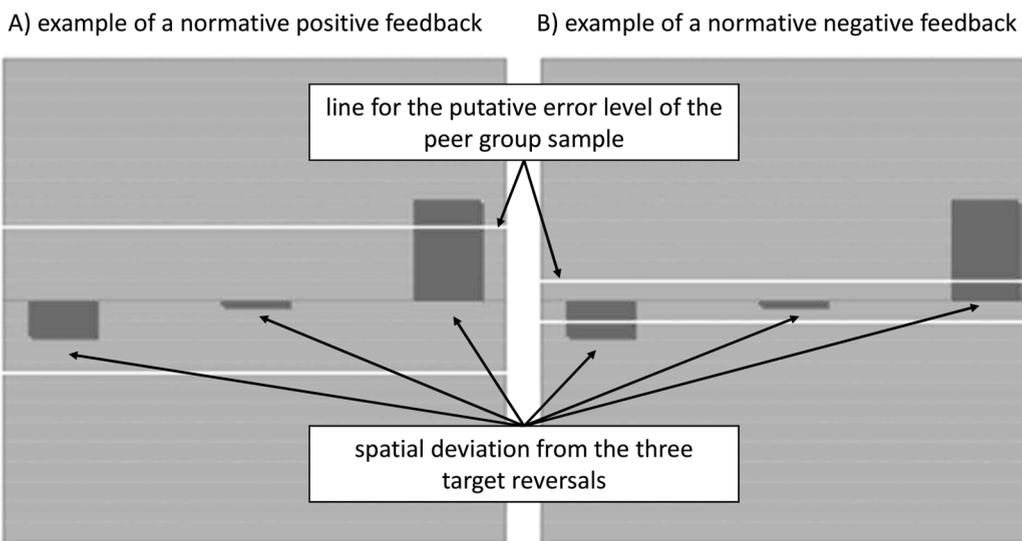


Fig. 3. Examples of feedback displays for normative positive (A) and normative negative feedback (B) showing the deviation for the three movement reversals with the putative error level of a peer-group sample.

As dependent variables, the median of *n-back errors* for the cognitive secondary task, and the mean *absolute error (AE)* as well as the mean *variable error (VE)* for movement precision and consistency were calculated. The factor *STDT* indicates the amount of dual-task costs, which stands for the degree of automaticity. The interaction *TIME* \times *STDT* then represents the automatization process. After completion of the *Pre-Test* and *Retention-Test*, participants rated their *perceived performance in relation to others (PPIRTO)* with a 5-point Likert scale item (manipulation check).

In addition to the *Pre-Test* and final *Retention-Test*, *Intermediate-Retention-Tests* (without feedback) with 12 single-task trials of the primary motor task were integrated prior to the feedback practice in sessions 1 to 5. These *Intermediate-Retention-Tests* were used to calculate the initial reference lines for the normative feedback manipulation.

2.4. Data analysis

We excluded trials with structural errors that had less or more than three reversals ($< 0.1\%$ of all test and practice trials were excluded). We did not exclude any trials with movement times above 1200 ms, as small deviations from the instructed maximum movement time are not critical for this experiment. Moreover, all participants followed the instruction, in consideration of the maximum movement time, to try to move rather rapidly.

The dependent measures for the lever-task are the *absolute error* for movement accuracy and the *variable error* for movement consistency (Schmidt & Lee, 2011). Both are measured in angular degrees. For the dependent measure *absolute error*, the mean absolute error for all movement reversals was calculated for each condition (single-task, dual-task). For the dependent measure *variable error*, the mean standard deviation was calculated for each of the three reversals. Then, the mean variable error of the reversals was calculated for each condition (single-task, dual-task). For the cognitive secondary task, the dependent variable is the *n-back error*. The *n-back error* is the median number of errors per trial (matches without response and false responses). According to the secondary task paradigm (Abernethy, 2001), the degree of automaticity is measured by analyzing the pre-post differences in dual-task costs (e.g., Blichke, 2000; Ruthruff et al., 2006).

To test *Hypothesis 1 (absolute error)* and *Hypothesis 2 (variable error)*, we used a repeated measures analysis of variance (ANOVA) with the factors *TIME (Pre-Test; Retention)* and *GROUP* on the single-task data.

To test *Hypothesis 3 (n-back error)*, we used a repeated measures ANOVA with the factors *TIME (Pre-Test; Retention)*, *STDT (Single-Task; Dual-Task)* and *GROUP*. To test for trade-off effects between the motor and the secondary task and to control for dual-task prioritization, we also analyzed the respective ANOVAs for the *absolute error* as well as the *variable error*. Any notable *TIME* \times *STDT* \times *GROUP* interaction effect would have to be considered for the discussion.

For a detailed analysis of the practice phase (see Appendix A), the *absolute error* and *variable error* were calculated for 15 practice blocks (early, middle and late block in 5 practice sessions). In session 1 with 114 trials, the block length was 38 trials, and in session 2–5 with 144 trials, block length was 48 trials. Practice performance was also tested for group effects with the 5 (*SESSION*) \times 3 (*BLOCK*) \times 2 (*GROUP*) ANOVAs. For all ANOVAs, the partial eta squared (η_p^2) was calculated as the effect size.

In addition, follow-up analyses were conducted with paired and unpaired *t*-tests. For all *t*-tests, Cohen's *d* for paired or unpaired samples was calculated as the effect size. A comparison of pre- and post-*PPIRTO*-values of the manipulation-check was calculated by a Wilcoxon-Test for each group. The alpha level was set to 0.05 for all statistical analyses.

3. Results

The analysis of the *Pre-Test* parallelization (3.1 Initial acquisition phase and *Pre-Test* comparisons) and dual-task prioritization (3.2 Prioritization Analysis) will be reported before the hypotheses related results are analyzed.

3.1. Initial acquisition phase and Pre-Test comparisons

The 3 (*GROUP*) \times 2 (*TIME: Pre-Acquisition; Pre-Test*) ANOVAs for the single-task performance indicate a significant main effect of *TIME* for the *absolute error*, $F(1, 39) = 4.74, p = .036, \eta_p^2 = .11$. The main effect of *TIME* for the *variable error* just falls short of significance, $F(1, 39) = 4.02, p = .052, \eta_p^2 = .09$. The *GROUP* \times *TIME* interaction is neither significant for the *absolute error*, $F(2, 39) = 0.58, p = .566, \eta_p^2 = .03$, nor the *variable error*, $F(2, 39) = 0.82, p = .921, \eta_p^2 < .01$. There are no significant *Pre-Test* differences between the groups for the *absolute error*, $F(2, 39) = 0.36, p = .698, \eta_p^2 = .02$ or the *variable error*, $F(2, 39) = 2.34, p = .110, \eta_p^2 = .11$, under single-task conditions.

Furthermore, the 3 (*GROUP*) \times 2 (*STDT*) ANOVAs for the *Pre-Test* do not indicate any significant *GROUP* \times *STDT* interactions, that is neither for the *n-back error*, $F(2, 39) = 0.49, p = .952, \eta_p^2 < .01$, nor the *absolute error*, $F(2, 39) = 0.13, p = .987, \eta_p^2 < .01$, or the *variable error*, $F(2, 39) = 1.76, p = .185, \eta_p^2 = .08$.

3.2. Prioritization analysis

The 2 (*TIME: Pre-Test; Retention*) \times 2 (*STDT*) \times 3 (*GROUP*) ANOVAs do not reveal a significant main effect of *STDT* for the *absolute error*, $F(1, 39) = 0.09, p = .773, \eta_p^2 < .01$, or the *variable error*, $F(1, 39) < 0.01, p = .942, \eta_p^2 < .01$. In contrast, the respective ANOVA for the *n-back error* shows a clear main effect of *STDT*, $F(1, 39) = 119.89, p < .001, \eta_p^2 = .76$. The considerable higher effect size (*n-back error*: $\eta_p^2 = .76$ vs. *absolute error*: $\eta_p^2 < .01$ and *variable error*: $\eta_p^2 < .01$) for dual-task costs for the *n-back* task is compatible to the instructed prioritization of the movement task during the dual-task test conditions. Accordingly, the automatization-

related effects will be tested by analysis of the n -back task. However, dual-task cost reductions for the *absolute error* and the *variable error* on the movement task side will be considered on account of the respective interaction terms in order to rule out misinterpretation owing to a dual-task cost trade off (see also Section 3.7 for details).

3.3. Testing the validity of the repeated dual-task test with regard to automaticity

The passive control group did not undergo a practice phase, but was tested in pre-test and retention test. This group can eliminate context-specific dual-task cost reductions from repeated dual-task testing (Agethen & Krause, 2012) as an alternative for single-task practice related dual-task cost reductions by automatization. Neither the 2 (*TIME: Pre-Test; Retention*) \times 2 (*STDT*) ANOVAs for the *absolute error*, $F(1, 13) < 0.01$, $p = .961$, $\eta_p^2 < .01$, nor the *n-back error*, $F(1, 13) = 0.32$, $p = .583$, $\eta_p^2 = .02$, calculated for this control group showed a significant two-way interaction. Thus, there is no indication for a reduction of dual-task costs caused by repeated testing (thereby replicating the results of Agethen & Krause, 2016). In addition, there is a main effect of *STDT* for the *n-back error*, $F(1, 13) = 37.56$, $p < .001$, $\eta_p^2 = .74$, but not for the *absolute error*, $F(1, 13) = 0.14$, $p = .909$, $\eta_p^2 < .01$, which again underscores successful prioritization of the primary motor task.

3.4. Normative feedback manipulation check – Perceived performance in relation to others (PPIRTO)

In the *Passive-Control-Group*, the *PPIRTO* did not change significantly (pre-value: $M = 3.07$, $SD = 0.62$; post-value: $M = 2.85$, $SD = 0.66$), $Z = -0.75$, $p = .454$. In the *Normative-Positive-Group*, the *PPIRTO* increased significantly to perceiving their performance as being above average (pre-value: $M = 2.86$, $SD = 0.66$; post-value: $M = 3.64$; $SD = 0.50$), $Z = -2.68$, $p = .008$. In the *Normative-Negative-Group*, the *PPIRTO* decreased significantly to perceiving their performance as being below average (pre-value: $M = 2.64$, $SD = 0.50$; post-value: $M = 2.36$; $SD = 0.63$), $Z = -2.00$, $p = .046$.

3.5. Learning – Movement accuracy (Hypothesis 1)

The 2 (*TIME: Pre-Test; Retention*) \times 2 (*GROUP: Normative-Positive-Group; Normative-Negative-Group*) ANOVA for the *absolute error* just fails to show a significant two-way interaction, $F(2, 39) = 3.88$, $p = .060$, $\eta_p^2 = .13$. Hypothesis 1 is not supported. There is a significant main effect of *TIME*, $F(1, 39) = 5.02$, $p = .031$, $\eta_p^2 = .11$.

The descriptive statistics and effect size, shows a large effect for *absolute error* reduction from *Pre-test* to *Retention* for the *Normative-Negative-Group*, $d = 0.91$, indicating a substantial increase in movement accuracy, while the *Normative-Positive-Group* only shows a small effect size, $d = 0.15$ (Fig. 4). Due to the failed significance of the two-way interaction, we did not calculate inferential statistics with respective tests for significance here.

3.6. Learning – Movement consistency (Hypothesis 2)

The 2 (*TIME: Pre-Test; Retention*) \times 2 (*GROUP: Normative-Positive-Group; Normative-Negative-Group*) ANOVA for the *variable error* indicates a significant two-way interaction, $F(1, 26) = 4.64$, $p = .041$, $\eta_p^2 = .15$. Follow up tests show significant reductions of the *variable error* for the *Normative-Positive-Group*, $t(13) = 3.25$, $p = .006$, $d = 0.87$, as well as for the *Normative-Negative-Group*, $t(13) = 5.62$, $p < .001$, $d = 1.50$, with a significantly higher reduction in the *Normative-Negative-Group*, $t(26) = -2.16$, $p = .041$, $d = -0.81$ (see Fig. 5). Hypothesis 2 is not supported.

3.7. Learning – Motor automaticity (Hypothesis 3)

The 2 (*TIME: Pre-Test; Retention*) \times 2 (*STDT*) \times 2 (*GROUP: Normative-Positive-Group; Normative-Negative-Group*) ANOVA for the *n-back error* indicates a significant three-way-interaction, $F(1, 26) = 9.27$, $p = .005$, $\eta_p^2 = .26$, as expected according to Hypothesis 3. In the follow-up tests (ANOVA *TIME* \times *STDT*), we find a significant *TIME* \times *STDT* interaction for the *Normative-Positive-Group*, $F(1,$

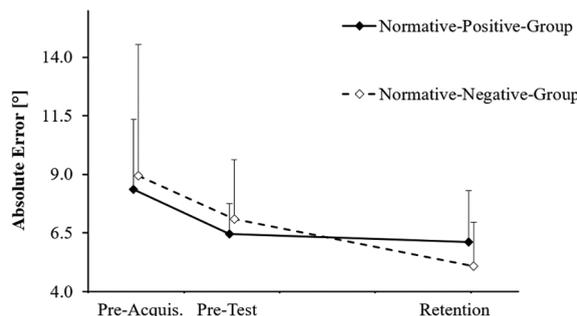


Fig. 4. Means and standard deviations of absolute errors in single-task conditions at Pre-Acquisition-Test, Pre-Test and Retention-Test.

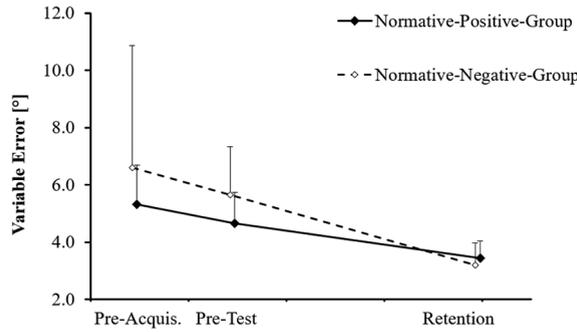


Fig. 5. Means and standard deviations of variable errors in single-task conditions at Pre-Acquisition-Test, Pre-Test and Retention-Test.

13) = 13.54, $p = .003$, $\eta_p^2 = .51$. In contrast, the *Normative-Negative-Group* clearly fails to show a *TIME* \times *STDT* interaction, $F(1, 13) = 0.66$, $p = .431$, $\eta_p^2 = .05$ (see Fig. 6). Further analyses for the *Normative-Positive-Group* indicate a significant reduction of the *n-back error* from *Pre-Test* to *Retention* in the dual-task condition, $t(13) = 4.81$, $p < .001$, $d = 1.29$, but not in the single-task-condition, $t(13) = 0.62$, $p = .547$, $d = 0.17$. According to these results, **Hypothesis 3** is supported.

In addition, we analyzed dual-task costs for the movement task to control for trade-off effects between the two tasks. The 2 (*TIME: Pre-Test; Retention*) \times 2 (*STDT: Single-Task; Dual-Task*) \times 2 (*GROUP*) ANOVA for the *absolute error* does not indicate a three-way interaction, $F(1, 26) = 0.02$, $p = .905$, $\eta_p^2 < .01$. Furthermore, there is no interaction of *STDT* \times *GROUP*, $F(1, 26) = 0.01$, $p = .932$, $\eta_p^2 < .01$. Likewise, the respective ANOVA for *variable error* does not indicate a three-way interaction, $F(1, 26) = 2.26$, $p = .145$, $\eta_p^2 = .08$ or interaction of *STDT* \times *GROUP*, $F(1, 26) = 0.72$, $p = .405$, $\eta_p^2 = .03$. Owing to the successful prioritization of the movement task (cf. Section 3.2) and the absence of any interactions with the factors *STDT* and *GROUP* for *absolute error* or *variable error* of the movement task, testing of **Hypothesis 3** is adequately related to the *n-back error* only, since there are no trade-off effects visible.

3.8. Learning – Intermediate Retention Tests

The 4 (*Intermediate Retention-Tests: IRI1 to IRI4*) \times 2 (*GROUP*) ANOVA for the *absolute error* in the intermediate retention tests shows significant main effects of *SESSION*, $F(2.40, 62.48) = 5.19$, $p = .005$, $\eta_p^2 = .17$, and *GROUP*, $F(1, 26) = 8.53$, $p = .007$, $\eta_p^2 = .25$. The *SESSION* \times *GROUP* interaction fails to be significant, $F(2.40, 62.48) = 0.70$, $p = .524$, $\eta_p^2 = .03$ (Fig. 7).

The 4 (*Intermediate-Retention-Test IRI1 to IRI4*) \times 2 (*GROUP*) ANOVA for the *variable error* in the intermediate retention tests shows significant main effects of *SESSION*, $F(2.47, 64.17) = 5.82$, $p = .003$, $\eta_p^2 = .18$, and *GROUP*, $F(1, 26) = 8.53$, $p = .007$, $\eta_p^2 = .25$. The *SESSION* \times *GROUP* interaction fails to be significant, $F(2.47, 64.17) = 0.94$, $p = .414$, $\eta_p^2 = .04$ (Fig. 8).

3.9. Movement time

In the *Pre-Test*, the movement time averages 1145 ms ($SD = 189$) under single-task conditions and 1101 ms ($SD = 208$) under dual-task conditions. In the *Retention-Test*, the movement time averages 1049 ms ($SD = 149$) under single-task conditions and 1011 ms ($SD = 168$) under dual-task conditions. The respective 2 (*TIME: Pre-Test; Retention*) \times 2 (*STDT*) \times 2 (*GROUP: Normative-*

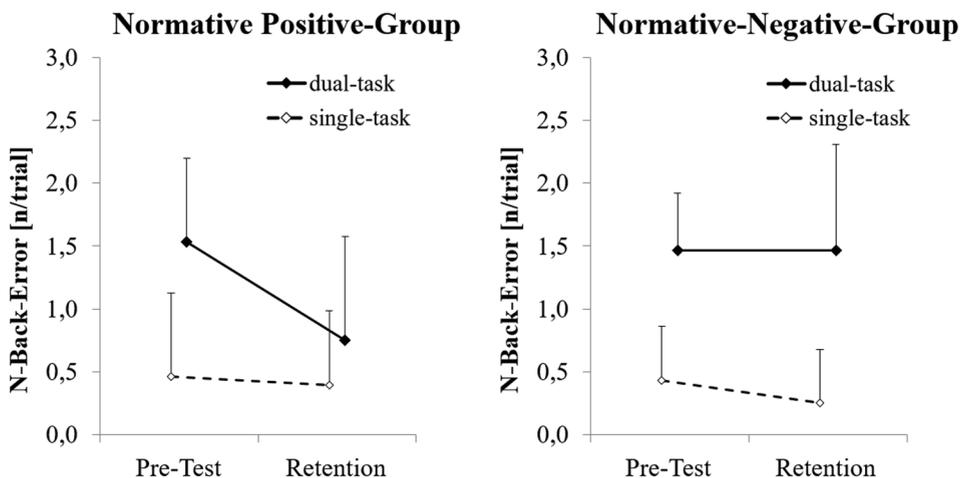


Fig. 6. Mean and Standard Deviation of *n-back errors* in single-task and dual-task conditions at Pre-Test and Retention-Test.

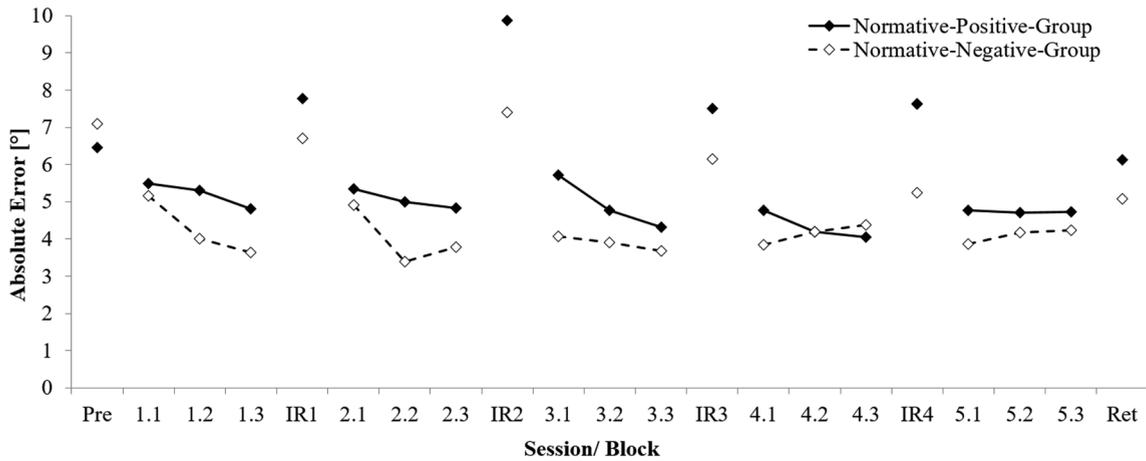


Fig. 7. Means of the absolute error in the practice blocks with feedback and tests without feedback (Pre-Test: Pre; intermediate retention tests: IR1-IR4; Retention-Test: Ret).

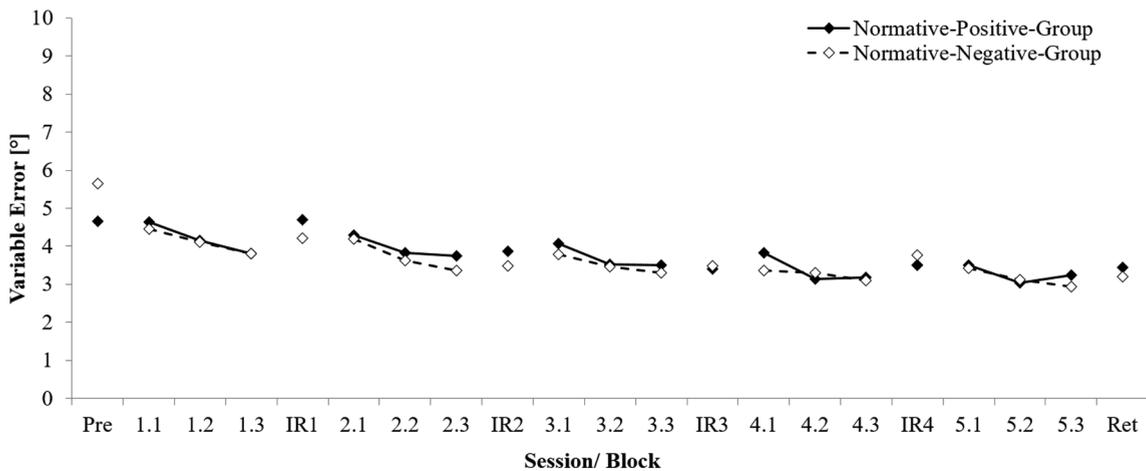


Fig. 8. Means of the variable error in the practice blocks with feedback and tests without feedback (Pre-Test: Pre; intermediate retention tests: IR1-IR4; Retention-Test: Ret).

Positive-Group; Normative-Negative-Group) ANOVA shows a main effect of TIME, $F(1, 26) = 6.46, p = .017, \eta_p^2 = .20$, as well as a main effect of STDT, $F(1, 26) = 9.39, p = .005, \eta_p^2 = .27$. There are no other significant main or interaction effects ($p \geq .376$).

During the practice sessions, all means are below the time restriction of 1200 ms on average. For a detailed analysis see Appendix A.

4. Discussion

4.1. Discussion on movement accuracy and movement consistency

Contrasting to earlier studies on normative feedback (Ávila et al., 2012; Lewthwaite & Wulf, 2010; Wulf et al., 2010) the feedback condition with more positive feedback valence did not show superior learning compared to the more negative feedback condition. Quite to the contrary, the Normative-Negative-Group reduced the absolute error even to a larger extend. Compared to the earlier studies mentioned, the amount of practice was substantially higher. A point of speculation might be that the participants in the Normative-Positive-Group did not believe in further improvements in the last practice sessions, as performance seemed to reach an asymptotic level. But, also the early practice phases and intermediate retention tests do not reveal beneficial learning for the normative positive conditions. The persisting feedback with positive valence indicated the participant’s performance to be at or above average and therefore may have restricted a further increase of the level of aspiration. This supposition also seems to be supported by the PPIRTO results. Hence, individuals are motivated to keep the level of effort and to maintain performance standard (Locke & Latham, 2002) as there is no incentive to invest more attentional resources to a goal that already has been achieved easily (Feedback Intervention Theory, Kluger & DeNisi, 1996).

In contrast, the Normative-Negative-Group had the putative information, that other participants were able to show a more accurate

performance. The information, that others were able to perform at a higher level, might have held up the belief in the reachability of higher performance levels and induced the setting of respective goals for learning. In terms of the goal-setting theory (Tzetzis, Kioumourtzoglou, & Mavromatis, 1997), this might have led to greater effort, directing attention on goal-relevant activities and development of strategies to improve performance. In certain circumstances, negative feelings like a moderate level of self-doubt after failure can motivate performers to exert more effort as has been shown in endurance tasks (Ede, Sullivan, & Feltz, 2017). Similar effects might occur in motor learning. Negative feedback also has been shown to motivate individuals more to concentrate on task achievement, while the effects of negative feedback might be moderated by situational and individual factors (Ilgen & Davis, 2000).

Our results are in line with several studies (Carter, Smith, & Ste-Marie, 2016; Ong, Lohse, & Hodges, 2015; Ong & Hodges, 2017; Patterson & Azizieh, 2012), who also failed to replicate findings of facilitation of learning through “success perceptions” (competency and success-related perceptions; Ong & Hodges, 2017) (e.g., Lewthwaite & Wulf, 2010). It is important to mention, that replications failed, even if the feedback manipulation were effective in affecting the perception of success (e.g., Ong & Hodges, 2017). Therefore, the general assumption of positive feedback as being more beneficial to learning, as it is assumed within some maybe over-simplified theoretical accounts (Wulf & Lewthwaite, 2016), cannot be maintained without looking at moderating factors and having a critical view on methods of earlier research. Several studies, which showed benefits of normative positive feedback on learning, did not integrate pre-tests in the experimental design. Thus, post-test differences in those studies might reflect pre-existing group differences, as already has been speculated by Ong and Hodges (2017). To this point, a recommendation for a general application of positive feedback cannot be derived, as empirical data is too heterogeneous. Theoretical and empirical approaches have to scrutinize in which temporal, situational and individual context normative positive feedback is effective for learning.

In perspective to individual factors, there is one important aspect to discuss for the *Normative-Positive-Group*, as there were putatively two performance level standards with diverging valences. One standard is related to the putative peer-group sample which was positive (in most of the feedback trials), whereas there is always a negative valence with regard to a perfect performance (deviation from zero line) which was also visible during each feedback presentation. The comparably high standard deviations of performance measures in the *Normative-Positive-Group* might reflect individual differences in orientation to one or the other of the two diverging performance level standards (see Fig. 7).

4.2. Discussion on motor automaticity

As expected, the normative feedback with positive valence facilitated the reduction of dual-task costs (in terms *n*-back error rates) more than normative feedback with negative valence, which even failed to reduce any dual-task costs at all. There are, however two different categories of errors in the *n*-back-task: omitted response on an *n*-back event (miss) and a response on a non-*n*-back event (false alarm; due to technical reasons, responses that are given delayed after the onset of the following stimulus are also categorized as false alarms). Additional detailed analyses for the type of errors in the *n*-back task revealed that there are more false alarms than misses in both groups. There are no group difference with respect to number of misses and false alarms in *Pre-Test*. Effect sizes for dual task cost reductions were large for false alarms, $\eta_p^2 = .51$, and for misses, $\eta_p^2 = .27$ in the *Normative-Positive-Group*, while these effect sizes were moderate in the *Normative-Negative-Group*, each $\eta_p^2 = .10$ (see Appendix B for more details on descriptive and inferential statistics). Thus, both experimental groups started with the same amount of errors in each of the two *n*-back task categories. And at the same time, group-specific changes in error rates due to feedback valence effects turned out pretty much the same for each of these error categories, thus corroborating the overall results pertaining to the differential group effect on motor automaticity.

The dual-task cost reductions in the *Normative-Positive-Group* are interpreted in terms of motor automatization, as the *Passive-Control-Group* does not show dual-task cost reductions due to repeated testing as an alternative explanation of dual-task cost reductions that may result from multiple test with a specific task combination (Agethen & Krause, 2012). This is in line with the EPAC-Hypothesis (Krause et al., 2018) and the assumption that negative feedback valence induces attentional processes like explicit hypothesis testing (Maxwell, Masters, Kerr, & Weedon, 2001), whereas normative positive feedback induces dopamine-dependent long-term potentiation of motor representations. According to the results on accuracy and consistency, this is not necessarily associated to a high level of performance, but to a well established representation in a motor code that enables motor control without the need to involve a high amount of attentional resources (Hikosaka et al., 1999). Also, progressive reductions of dual-task costs (i.e. increases in automaticity) are typically found even after single-task performance measures have already reached an asymptotic level (Puttemans, Wenderoth, & Swinnen, 2005). Regarding automaticity, the results are also consistent with an assumption of Lewthwaite and Wulf (2010) who found higher frequencies of movement corrections in a balance task with positive normative feedback and discussed this in terms of a higher degree of motor automaticity. However, associations between high frequent movement corrections and automatic control were strictly speculative and not based on empirical data in their study.

The main finding of a higher reduction in dual-task costs in the current experiment's *Normative-Positive-Group* complements earlier findings on feedback valence and motor automatization, where a bandwidth approach was used to manipulate feedback valence using the same motor task (Agethen & Krause, 2016). Here and also in a similar work on feedback frequency (Krause et al., 2018), feedback effects on automaticity were also dissociated from motor accuracy and consistency, as low error-feedback frequency (Krause et al., 2018) and a bandwidth feedback condition (Agethen & Krause, 2016) facilitated automaticity compared to permanent error feedback without making a difference in terms of accuracy and consistency. In a closely related research paradigm on varied success-rates during motor practice (manipulated by task difficulty; e.g., varied distance or target size in golf putting or underhand throwing), benefits for practice conditions with higher success-rates are evident in terms of automaticity (i.e., lower dual-task costs) going along with benefits in single-task performance in some cases (e.g., Capio, Poolton, Sit, Holmstrom, & Masters, 2013; Masters et al., 2001; Savelsbergh, Canal-Bruland, & van der Kamp, 2012), while they sometimes benefit automaticity without affecting single-task

performance differently (e.g., Lam, Maxwell, & Masters, 2010; Poolton, Masters, & Maxwell, 2007). Due to differences in tasks and procedures (e.g., pre-post design vs. post-design; early vs. late retention tests) discussing moderating variables would be highly speculative.

Increased automaticity, which is not associated with better single-task performance regarding accuracy is even more distinct in the current experiment as single-task performance in the normative positive feedback condition is even negatively affected compared to the feedback condition with more negative valence. Consequently, this dissociative effect of feedback valence on accuracy and automaticity calls for replication with a priori stated hypotheses.

One important difference between our and other studies is the frequency of performance related information presentation. For example, the errorless learning paradigm usually uses tasks where information on task performance is present on each trial. In the current study, we provided feedback on only 14% of the trials. We decided to use 14% as this frequency allows automatization to occur, while 100% in this experimental setting does not seem to allow any automatization at all (Agethen & Krause, 2016; Krause et al., 2018). Undoubtedly, feedback frequency by itself is a moderating variable on motor automatization as well as on single-task learning effects.

At this point, we have gained some basic knowledge about the effects of feedback valence on motor automaticity and even first indications that automatization and accuracy learning develop independently moderated by feedback valence, other possible situational and individual factors notwithstanding. From an applied perspective, the feedback valence dependent dissociation raises questions on how to combine the beneficial effects of positive and negative valence properly over the course of practice. The adequate scheduling of feedback valence throughout practice blocks and sessions should be addressed in future studies in order to develop sophisticated complex feedback strategies to achieve both, a high level of accuracy and automatization in the long term.

As both the primary and the secondary task in our present study to some extent share motor response demands (insofar as in both cases the upper extremities are involved), there might be residual dual-task costs (DTC) outlasting even extensive practice. However, this case seems to be quite unlikely in the light of the results of a previous study incorporating a combination of primary and secondary tasks identical to the present one (Agethen & Krause, 2016). In that study, dual-task costs were reduced even near to zero, leaving no room for any relevant residual DTC-effect. Moreover, even if there actually were any dual-task costs due to “low level” structural interference (cf. Heuer & Wing, 1984, and more recently, Klapp, Maslovat, & Jagacinski, 2018), their amount ought to be the same for both of our experimental groups, and at the same time they should be unrelated to any dual-task costs resulting from mental work-load exceeding central processing resources. Thus, only changes in those latter, cognitively demanding processes according to our point of view could be causally related to the treatment-induced performance changes found in our present study.

5. Conclusion

Consistent with earlier studies, motor automaticity has been shown to be moderated by feedback manipulation. Inducing the impression to be better than others by normative feedback facilitated motor automaticity. The effect on automaticity is dissociated from the effect on motor performance in terms of accuracy. Inconsistent with earlier studies, inducing the impression to be worse than others facilitated motor accuracy. Future studies might reveal if the inconsistent findings might be explainable by moderating variables as the category of the motor task or amount of practice.

Appendix A

See [Table A.1](#)

Table A.1

Main effects and interactions with factor GROUP of GROUP \times SESSION \times BLOCK ANOVAs.

	<i>F</i> (<i>df</i>)	<i>p</i>	η^2_p
<i>Absolute error</i>			
GROUP	3.88 (1, 26)	.060	.13
SESSION	0.96 (2.89, 75.04)	.432	.04
BLOCK	7.30 (1.17, 30.31)	.009*	.22
GROUP \times SESSION	0.91 (2.89, 75.04)	.438	.03
GROUP \times BLOCK	0.42 (1.17, 30.31)	.550	.02
GROUP \times SESSION \times BLOCK	2.47 (4.22, 109.68)	.046* ¹	.09
<i>Variable error</i>			
GROUP	1.23 (1, 26)	.277	.05
SESSION	44.66 (3.09, 80.23)	< .001*	.63
BLOCK	49.97 (1.59, 41.33)	< .001*	.62
GROUP \times SESSION	0.27 (3.09, 80.23)	.855	.01
GROUP \times BLOCK	1.72 (1.59, 41.33)	.196	.06
GROUP \times SESSION \times BLOCK	1.04 (5.53, 143.79)	.396	.04
<i>Movement time</i> ²			
GROUP	< 0.01 (1, 26)	.962	< .01
SESSION	0.39 (1.79, 46.67)	.655	.02

(continued on next page)

Table A.1 (continued)

	<i>F(df)</i>	<i>p</i>	η^2_p
BLOCK	1.12 (1.69, 43,86)	.327	.04
GROUP × SESSION	0.35 (1.79, 46,67)	.682	.01
GROUP × BLOCK	0.70 (1.79, 46,67)	.481	.03
GROUP × SESSION × BLOCK	0.68 (2.91, 75.53)	.562	.03

¹ Post-hoc follow-up single comparisons reveal significant group differences in session 1 (late block), $t(26) = 2.42$; $p = .023$, session 2 (middle block), $t(15.04) = 3.51$, $p = .003$, and session 3 (early block), $t(19.16) = 3.36$, $p = .003$.

² During the practice sessions, the movement time averages 1066 (Practice 1; $SD = 166$); 1050 (Practice 2; $SD = 163$); 1036 (Practice 3; $SD = 156$); 1060 (Practice 4; $SD = 157$); 1058 (Practice 5; $SD = 156$). The 5 (SESSION) × 3 (BLOCK) × 2 (GROUP) ANOVA of the practice trials for the movement time does not reveal a main effect of GROUP, $F(1, 26) < 0.01$, $p = .962$, $\eta^2_p < .01$, or interaction with the factor GROUP (all with $\eta^2_p < .03$). All means are below the time restriction of 1200 ms on average, but the movement time limit was exceeded in 18.0% ($SD = 24.9\%$) of the trials in the Normative-Negative-Group and in 19.5% ($SD = 22.6\%$) in the Normative-Positive-Group. The percentage of trials with exceeded movement time did not differ, $t(26) = 0.168$; $p = .868$; $d = 0.06$.

Appendix B

See Tables B.1 and B.2

Table B.1

Descriptive statistics (means and standard deviations in parentheses) and inferential statistics for misses in the *n*-back task.

	Descriptives		<i>ST vs. DT</i>		<i>TIME × STDT</i>	
	ST	DT	<i>p</i>	<i>d</i>	<i>p</i>	η^2_p
<i>Pre-Test</i> ¹						
Normative-Negative-Group	0.11 (0.29)	1.14 (0.46)	< 0.001	2.08	–	–
Normative-Positive-Group	0.25 (0.55)	1.04 (0.63)	0.001	1.22	–	–
<i>Retention</i>						
Normative-Negative-Group	0.14 (0.31)	0.96 (0.66)	< 0.001	1.43	.254	.10
Normative-Positive-Group	0.32 (0.50)	0.64 (0.66)	0.045	0.59	.048	.27

¹ There are no pre-test differences between the groups for the misses (ST: $p = .395$, $d = 0.33$ or DT: $p = .612$, $d = -0.19$).

Table B.2

Descriptive statistics (means and standard deviations in parentheses) and inferential statistics for false alarms in the *n*-back task.

	Descriptives		<i>ST vs. DT</i>		<i>TIME × STDT</i>	
	ST	DT	<i>p</i>	<i>d</i>	<i>p</i>	η^2_p
<i>Pre-Test</i> ¹						
Normative-Negative-Group	0.46 (0.50)	1.46 (0.46)	< .001	0.94	–	–
Normative-Positive-Group	0.50 (0.76)	1.57 (0.70)	< .001	1.47	–	–
<i>Retention</i>						
Normative-Negative-Group	0.21 (0.32)	1.46 (0.84)	< .001	1.70	.265	.10
Normative-Positive-Group	0.43 (0.83)	0.79 (0.80)	.027	0.67	.003	.51

¹ There are no pre-test differences between the groups for the false alarms (ST: $p = .884$; $d = 0.06$; DT: $p = .637$, $d = 0.18$).

References

- Abernethy, B. (2001). Attention. In R. N. Singer, H. A. Hausenblas, & C. M. Janelle (Eds.). *Handbook of Sport Psychology* (pp. 53–85). New York: Wiley.
- Agethen, M., & Krause, D. (2012). Reduced dual task interference in multiple repeated dual-task tests: Automatization or task integration? In G. Juras, & K. Slomka (Eds.). *Current research in motor control IV – From Theory to Implementation* (pp. 8–14). Katowice: AWF Katowice.
- Agethen, M., & Krause, D. (2016). Effects of bandwidth feedback on the automatization of an arm movement sequence. *Human Movement Science*, 45, 71–83. <https://doi.org/10.1016/j.humov.2015.11.005>.
- Ávila, L. T., Chiviawowsky, S., Wulf, G., & Lewthwaite, R. (2012). Positive social-comparative feedback enhances motor learning in children. *Psychology of Sport and Exercise*, 13, 849–853.
- Blichke, K. (2000). Two procedures, one mechanism? Recent findings on the automation of voluntary movements. *Journal of Human Kinetics*, 4, 3–16.
- Blichke, K., Wagner, F., Zehren, B., & Brueckner, S. (2010). Dual-task practice of temporally structured movement sequences augments integrated task processing, but not automatization. *Journal of Human Kinetics*, 25, 5–15. <https://doi.org/10.2478/v10078-010-0026-1>.
- Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective, & Behavioral Neuroscience*, 7, 356–366. <https://doi.org/10.3758/CABN.7.4.356>.
- Büsch, D., Hagemann, N., & Bender, N. (2010). The dimensionality of the Edinburgh Handedness Inventory: An analysis with models of the item response theory.

- Laterality*, 15, 610–628. <https://doi.org/10.1080/13576500903081806>.
- Capio, C. M., Poolton, J. M., Sit, C. H. P., Holmstrom, M., & Masters, R. S. W. (2013). Reducing error benefits the field-based learning of a fundamental movement skill in children. *Scandinavian Journal for Medicine and Science in Sports*, 23, 181–188.
- Carter, M. J., Smith, V., & Ste-Marie, D. M. (2016). Judgments of learning are significantly higher following feedback on relatively good versus relatively poor trials despite no actual learning differences. *Human Movement Science*, 45, 63–70. <https://doi.org/10.1016/j.humov.2015.11.006>.
- Chein, J. M., & Schneider, W. (2012). The brain's learning and control architecture. *Current Directions in Psychological Science*, 21, 78–84. <https://doi.org/10.1177/0963721411434977>.
- Ede, A., Sullivan, P. J., & Feltz, D. L. (2017). Self-doubt: Uncertainty as a motivating factor on effort in an exercise endurance task. *Psychology of Sport and Exercise*, 28, 31–36. <https://doi.org/10.1016/j.psychsport.2016.10.002>.
- Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proceedings of the National Academy of Sciences*, 108(Supplement 3), 15647–15654. <https://doi.org/10.1073/pnas.1014269108>.
- Heckhausen, J., & Heckhausen, H. (2018). *Motivation and Action*. Champaign: Springer 10.1007/978-3-642-75961-1.
- Heuer, H., & Wing, A. M. (1984). Doing two things at once: Process limitations and interactions. In I. Smyth, & A. Wing (Eds.). *The Psychology of Human Movement* (pp. 183–213). London: Academic Press.
- Hikosaka, O., & Isoda, M. (2010). Switching from automatic to controlled behavior: Cortico-basal ganglia mechanisms. *Trends in Cognitive Sciences*, 14, 154–161. <https://doi.org/10.1016/j.tics.2010.01.006>.
- Hikosaka, O., Nakahara, H., Rand, M. K., Sakai, K., Lu, X., Nakamura, K., ... Doya, K. (1999). Parallel neural networks for learning sequential procedures. *Trends in Neurosciences*, 22, 464–471. [https://doi.org/10.1016/S0166-2236\(99\)01439-3](https://doi.org/10.1016/S0166-2236(99)01439-3).
- Hikosaka, O., Nakamura, K., Sakai, K., & Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current opinion in neurobiology*, 12, 217–222. [https://doi.org/10.1016/S0959-4388\(02\)00307-0](https://doi.org/10.1016/S0959-4388(02)00307-0).
- Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109, 679–709. <https://doi.org/10.1037//0033-295X.109.4.679>.
- Ilgel, D., & Davis, C. (2000). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology*, 49, 550–565. <https://doi.org/10.1111/1464-0597.00031>.
- Klapp, S. T., Maslovat, D., & Jagacinski, R. J. (2018). The bottleneck of the psychological refractory period effect involves timing of response initiation rather than response selection. *Psychonomic Bulletin & Review*, 24, 1–19. <https://doi.org/10.3758/s13423-018-1498-6>.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254–284. <https://doi.org/10.1037/0033-2909.119.2.254>.
- Krause, D., Agethen, M., & Zobe, C. (2018). Error feedback frequency affects automaticity but not accuracy and consistency after extensive motor skill practice. *Journal of Motor Behavior*, 50, 144–154. <https://doi.org/10.1080/00222895.2017.1327406>.
- Lam, W. K., Maxwell, J. P., & Masters, R. S. W. (2010). Probing the allocation of attention in implicit (motor) learning. *Journal of Sports Sciences*, 28, 1543–1554. <https://doi.org/10.1080/02640414.2010.517543>.
- Lewthwaite, R., & Wulf, G. (2010). Social-comparative feedback affects motor skill learning. *The Quarterly Journal of Experimental Psychology*, 63, 738–749. <https://doi.org/10.1080/17470210903111839>.
- Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57, 705–714.
- Maxwell, J. P., Masters, R. S. W., Kerr, E., & Weedon, E. (2001). The implicit benefit of learning without errors. *The Quarterly Journal of Experimental Psychology Section A*, 54, 1049–1068. <https://doi.org/10.1080/1713756014>.
- Ong, N. T., & Hodges, N. J. (2017). Balancing our perceptions of the efficacy of success-based feedback manipulations on motor learning. *Journal of Motor Behavior*. Advance online Publication. <https://doi.org/10.1080/00222895.2017.1383227>.
- Ong, N. T., Lohse, K. R., & Hodges, N. J. (2015). Manipulating target size influences perceptions of success when learning a dart-throwing skill but does not impact retention. *Frontiers in Psychology*, 6, 1378. <https://doi.org/10.3389/fpsyg.2015.01378>.
- Patterson, J. T., & Azizieh, J. (2012). Knowing the good from the bad: Does being aware of KR content matter? *Human Movement Science*, 31, 1449–1458. <https://doi.org/10.1016/j.humov.2012.04.004>.
- Poolton, J. M., Masters, R. S. W., & Maxwell, J. P. (2007). Passing thoughts on the evolutionary stability of implicit motor behaviour: Performance retention under Physiological Fatigue. *Consciousness and Cognition*, 16, 456–468. <https://doi.org/10.1016/j.concog.2006.06.008>.
- Puttemans, V., Wenderoth, N., & Swinnen, S. P. (2005). Changes in brain activation during the acquisition of a multifrequency bimanual coordination task: From the cognitive stage to advanced levels of automaticity. *Journal of Neuroscience*, 25, 4270–4278. <https://doi.org/10.1523/JNEUROSCI.3866-04.2005>.
- Ruthruff, E., Van Sleest, M., Johnston, J. C., & Remington, R. (2006). How does practice reduce dual-task interference: Integration, automatization, or just stage shortening? *Psychological Research*, 70, 125–142. <https://doi.org/10.1007/s00426-004-0192-7>.
- Savelsbergh, G., Canal-Bruland, R., & van der Kamp, J. (2012). Error reduction during practice: A novel method for learning to kick free-kicks in soccer. *International Journal of Sports Science & Coaching*, 7, 47–56. <https://doi.org/10.1260/1747-9541.7.1.47>.
- Schmidt, R. A., & Lee, T. D. (2011). *Motor Control and Learning: A Behavioral Emphasis* (5th ed.). Champaign: Human Kinetics.
- Seidler, R. D., Kwak, Y., Fling, B. W., & Bernard, J. A. (2013). Neurocognitive mechanisms of error-based motor learning. In M. J. Richardson, M. A. Riley, & K. Shockley (Eds.). *Advances in Experimental Medicine and Biology. Progress in Motor Control* (pp. 39–60). New York: Springer. https://doi.org/10.1007/978-1-4614-5465-6_3.
- Tzetzis, G., Kioumourtzoglou, E., & Mavromatis, G. (1997). Goal setting and feedback for the development of instructional strategies. *Perceptual and motor skills*, 84(3 Supplement), 1411–1427. <https://doi.org/10.2466/pms.1997.84.3c.1411>.
- Wulf, G., & Lewthwaite, R. (2016). Optimizing performance through intrinsic motivation and attention for learning: The OPTIMAL theory of motor learning. *Psychonomic Bulletin & Review*, 23, 1382–1414. <https://doi.org/10.3758/s13423-015-0999-9>.
- Wulf, G., Chiviacowsky, S., & Lewthwaite, R. (2010). Normative feedback effects on learning a timing task. *Research Quarterly for Exercise and Sport*, 81, 425–431. <https://doi.org/10.1080/02701367.2010.10599703>.