RESEARCH ARTICLE | *Control of Movement*

The effect of instruction on motor skill learning

Nicola J. Popp,[1] Atsushi Yokoi,[2,3] Paul L. Gribble,[1,4,5,6] **and** Jörn Diedrichsen[1,7,8]

[1]*The Brain and Mind Institute, University of Western Ontario, London, Ontario, Canada;* [2]*Graduate School of Frontier Biosciences, Osaka University, Osaka, Japan;* [3]*Center for Information and Neural Networks (CiNet), NICT, Osaka, Japan;* [4]*Department of Psychology, University of Western Ontario, London, Ontario, Canada;* [5]*Department of Physiology & Pharmacology, Schulich School of Medicine & Dentistry, University of Western Ontario, London, Ontario, Canada;* [6]*Haskins Laboratories, New Haven, Connecticut;* [7]*Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario, Canada; and* [8]*Department of Computer Science, University of Western Ontario, London, Ontario, Canada*

Popp NJ, Yokoi A, Gribble PL, Diedrichsen J. The effect of instruction on motor skill learning. *J Neurophysiol* 124: 1449–1457, 2020. First published September 30, 2020; doi:10.1152/jn.00271.2020.—Many motor skills are learned with the help of instructions. In the context of complex motor sequences, instructions often break down the movement into chunks that can then be practiced in isolation. Thus, instructions shape an initial cognitive representation of the skill, which in turn guides practice. Are there ways of breaking up a motor sequence that are better than others? If participants are instructed in a way that hinders performance, how much practice does it take to overcome the influence of the instruction? To answer these questions, we used a paradigm in which participants were asked to perform finger sequences as fast and accurately as possible on a keyboard-like device. In the initial phases of training, participants had to explicitly remember and practice two- or three-digit chunks. These chunks were then combined to form seven 11-digit sequences that participants practiced for the remainder of the study. Each sequence was broken up into chunks in a way such that the instruction was either aligned or misaligned with the basic execution-level constraints. We found that misaligned chunk instruction led to an initial performance deficit compared with the aligned chunk instruction. Overall, instructions still influenced the temporal pattern of performance after 10 days of subsequent training, with shorter interpress intervals within a chunk compared with between chunks. However, for the misaligned instructed sequences, this temporal pattern was altered more rapidly, such that participants could overcome the induced performance deficit in the last week. At the end of training, participants found idiosyncratic, but interindividually stable, ways of performing each sequence.

NEW & NOTEWORTHY Instructions often break down motor sequences into smaller parts, such that they can be more easily remembered. Here, we show that different ways of breaking down a finger sequence can subsequently lead to better or worse performance. The initial instruction still influenced the temporal performance pattern after 10 days of practice. The results demonstrate that the initial cognitive representation of a motor skill strongly influences how a skill is learned and performed.

chunking; habits; instructions; motor learning; motor skill

INTRODUCTION

Many motor skills are learned with the help of instructions, be it from peers, parents, teachers, or videos. Such instructions often break down longer, complicated sequences of actions, for example, tying one's shoelaces, into smaller elements that can be more easily remembered and practiced. Thus, instructions help to build up an initial cognitive representation of the motor skill, which in turn guides physical practice (Green and Flowers 1991; Hodges and Franks 2002; Masters 1992; Meier et al. 2020; Wulf et al. 1998). However, it is unclear for how long these initial cognitive representations influence motor behavior. At what point does the learner find an optimal way of performing the action independent of the initial instruction? Furthermore, are some ways of building an initial cognitive representation of a motor sequence better than others?

We investigated these questions in the context of a discrete sequence production (DSP) task, in which participants perform a series of single finger presses as fast as possible while having full knowledge of the sequence (Abrahamse et al. 2013; Verwey 2001). Learning in this task depends on both cognitive and motor processes (Diedrichsen and Kornysheva 2015; Wong et al. 2015). Initial performance relies strongly on forming a declarative memory of the sequence (de Kleine and Verwey 2009; Verwey and Dronkert 1996; Verwey et al. 2009, 2010). This initial declarative memory (or cognitive representation) of the motor sequence is often characterized by chunking—the process of breaking down a (to make it: breaking down a long sequence of items) long sequence of items into smaller subsets, which has been shown to aid memorization (Halford et al. 1998; Miller 1956; Solopchuk et al. 2016). In the context of movement sequences, this chunked memory representation has been shown to influence the actual motor performance, with participants inserting larger pauses between chunks than between finger presses within a chunk (Verwey 1996; Verwey and Dronkert 1996). This cognitively induced chunking pattern (de Kleine and Verwey 2009; Verwey and Dronkert 1996; Verwey et al. 2009, 2010) has been found to remain stable over the course of a few days of practice (Verwey and Eikelboom 2003).

We hypothesized that the way the initial declarative memory of a motor sequence is chunked will either facilitate or impede subsequent skill learning. To test this idea, we instructed

Correspondence: J. Diedrichsen (jdiedric@uwo.ca).

participants to memorize long sequences of finger presses by first practicing smaller two- to three-digit "chunks" on an isometric keyboard-like device. Participants were then trained on the seven 11-digit sequences that were made up of the two- to three-digit chunks that they had learned before. Each sequence was subdivided and instructed in two different ways: in a counterbalanced within-subject design, half of the sequences for each participant were instructed using an aligned chunking structure, in which the boundaries between chunks were aligned with finger transitions that were difficult to execute. These transitions constituted natural breaking points, and we hypothesized that the participants should be able to use the time needed to execute these transitions to recall the next chunk. In addition, we kept easy finger transitions, such as runs (e.g., 123), together within a chunk, enabling participants to execute these quickly. The other half of the sequences were instructed using a misaligned chunking structure, in which these easy finger transitions were artificially broken up by chunk boundaries, and difficult transitions were kept within a chunk. We hypothesize that sequences that were learned under a misaligned structure would be executed more slowly. After the instruction phase, participants practiced the sequences over the course of 3 wk, allowing us to investigate the influence of the initial instruction on subsequent motor skill learning.

Specifically, we investigated three questions: first, do the initial chunk instructions lead to stable movement patterns that persist over 3 wk of training? Second, can different chunk instructions on the exact same sequence lead to movement patterns that are superior or inferior in terms of performance? We hypothesized that sequences learned using the misaligned instruction would be performed slower compared with sequences learned with the aligned instruction. Third, if we can induce deficits in performance using suboptimal chunk instructions, for how long are these maladaptive patterns maintained?

METHODS

Participants

In total, 40 participants who reported no neurological conditions were recruited for the study (30 females; age: 19–33). Thirty-two of them were randomly assigned to learn the sequences with one of the two chunk sets (Fig. 1). Eight additional participants were recruited as a control group that did not receive any chunk instructions. All participants were right-handed based on the Edinburgh Handedness Inventory and completed informed consent. On average, participants had received 4.68 (±5.55) years of musical training, with 55% reported having more than 6 mo of experience playing the piano. Although participants with piano experience performed the sequences faster than participants with no experience and the number of practice years correlated with execution speed (movement time, MT), the number of participants' prior musical experience did not have a qualitative influence on participants' chunking behavior. The study protocol was approved by the ethics board of the University of Western Ontario.

Apparatus

A custom-built five-finger keyboard was used (Fig. 1A). The keys were not depressible but were equipped with a force transducer (FSG-15N1A, Sensing and Control, Honeywell) underneath each key that measured participants' isometric force production with a repeatability of <0.02 N and a dynamic range of 16 N (Wiestler and Diedrichsen 2013; Wiestler et al. 2014; Yokoi et al. 2017). The measured force at each key was digitally sampled at 200 Hz.
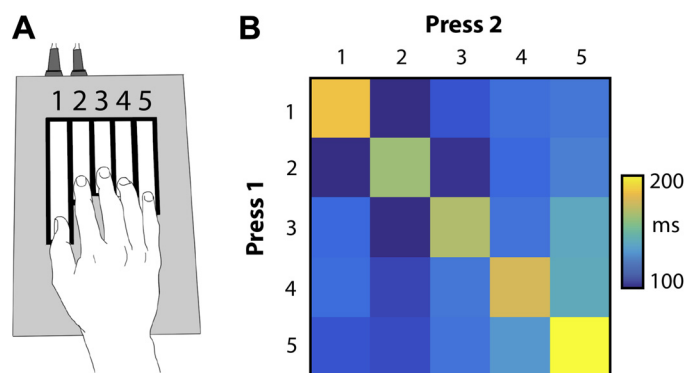

Fig. 1. Apparatus and two-finger transition execution speed. A: isometric keyboard-like device. Each key was associated with a number (these numbers were not shown to the participants but verbally explained). B: data from the independent baseline study in which participants performed all possible combinations of two-digit transitions. Matrix indicates the median interpress interval (IPI) to produce the transition between pairs of keypresses. Indicated values are means over $n = 7$ participants.

Discrete Sequence Production Task

We used a discrete sequence production (DSP) task in which participants executed sequences of 2, 3, or 11 keypresses as fast as possible while keeping their error rate under 15%. Each finger was associated with a number (thumb = 1, index = 2, middle = 3, ring = 4, and pinky = 5; Fig. 1A). Each trial started with the visual presentation of the sequence to be executed and was completed once the participants pressed the amount of presented numbers.

A keypress was registered when the measured force first exceeded 1.5 N. A key release was marked when the force measured at the same key first fell below 1 N. The magnitude of the force applied to each key was represented by five lines on an LCD monitor, with the height of the line representing the force at the corresponding key. No pause between presses was required, and thus, some coarticulation between fingers emerged with faster execution. However, to prevent participants from pressing several keys at once, the previously pressed key had to be released before a new key could be registered as pressed.

Immediately after the keypress threshold was reached, participants received visual and auditory feedback. If the correct key was pressed, the color of the corresponding digit changed from white to green and a sound was presented (same sound for each key). If the incorrect key was pressed, the cue turned red and a lower-pitch sound was presented.

After each trial, participants received points based on their accuracy (whether all presses in the sequence were correct) and movement time (MT; the time between the first keypress and last key release). Correct sequences that were performed faster than the MT threshold were rewarded with 1 point. MTs that were 20% faster than the threshold were rewarded with 3 points. Incorrect presses or MTs exceeding the threshold resulted in 0 points. At the end of each block, participants received feedback on their error rate, median MT, points obtained during the block, and total points obtained during the session. To motivate participants to continue to improve their performance, we adjusted the MT threshold by lowering it by 500 ms after each block in which the participants performed with an error rate of 15% or lower and had a median MT faster than the current threshold. This manipulation resulted in a stable overall error rate of 14.6% (SD = 2.6%). In 27% of the trials, participants received 1 point, and in 34% of the trials, they received 3 points.

Baseline Study for Measuring Execution-Level Constraints

One of the aims of the study was to design specific ways of chunking a sequence that would induce either better or worse performance. We hypothesized that it would be advantageous to have chunk boundaries

fall on transitions between fingers that are, based on execution-level constraints, executed slowly. We define execution-level constraints as factors arising from the neural control of movement, biomechanics, and characteristics of the keyboard device, independent of cognitive factors. To determine these constraints on finger transition speed, i.e., how fast participants can naturally execute each of the 25 possible two-finger transitions (e.g., 12, 13, 25), we recruited seven participants (five females, age = 21–27) for a 3-day study. None of the participants in this experiment participated in the later main experiment. Participants executed all possible two-finger transitions (e.g., 25) and three-finger transitions (e.g., 125), each eight times per day. Each sequence was presented twice in a row. Each day, participants completed eight blocks with 150 trials each. The setup, task, and feedback were the same as in the main experiment.

Because participants only had to plan and execute two or three finger presses, we assumed that cognitive constraints, chunking, or planning processes should not have limited performance. Thus, the average speed of these transitions can be taken as a characterization of the execution-level constraints of our specific task. The data from the two-finger transitions revealed a clear pattern (Fig. 1B), in which transitions between adjacent fingers (e.g., 12, 23, 32) could be executed on average 68.5 ms faster than finger repetitions (e.g., 55, 33, 22). We tested this difference by comparing the average speed of adjacent finger presses with the average speed of repetitions with a paired t test [$t(6) = 13.965$, $P = 8.404\ e{-}06$; Fig. 1B]. To press the same finger twice, the force applied to the key had to first exceed the press threshold, then go below the release threshold and then cross the press threshold again. This rapid alternation of forces takes time to produce. In contrast, for two adjacent fingers, the second finger press can be initiated (have already reached the press threshold but have not yet been registered) before the previous finger is released, making it easier to rapidly produce this force pattern.

The overall 5×5 pattern of interpress intervals (IPIs) was stable across participants (average correlation $r = 0.689$) and days ($r = 0.894$), even though participants improved their overall speed from 157 ms on the first day to 114 ms on the third day. The same pattern was also apparent for the three-finger transition data. If we broke up the three-finger transitions into the constituent IPIs, the average pattern correlated with the two-finger transition data with $r = 0.913$.

Experimental Design

To experimentally impose a particular way of chunking, we instructed participants in the experimental group to memorize and perform a set of two to three keypress chunks (Fig. 2A). These chunks were later combined to form the training sequences (Fig. 2B). Our goal was to impose beneficial or detrimental motor patterns on participants' performance. For this, we used the finding from the execution-level constraint baseline study that finger repetitions are performed slower than presses of adjacent fingers. We designed sequences such that they would include both fast transitions (runs, e.g., 123—three-digit transition with two adjacent finger transitions that are either descending or ascending) and slow finger repetitions (e.g., 113). In the "aligned" chunk structure, we inserted chunk boundaries such that they fell on difficult finger transitions, which were executed slowly. We hypothesized that participants could use the time required to perform these difficult finger transitions to recall the next chunk, which should benefit the overall performance. In this chunk structure, the three-digit "runs" (i.e., 123) were also kept intact within a chunk. We predicted that learning the sequence using this chunk structure would be beneficial to performance speed (Fig. 2C). In the misaligned chunk structure, we placed chunk boundaries in a way that divided up fast finger transitions such as runs (e.g., 123), thereby breaking up parts of the sequence that could otherwise be performed very quickly. Adding chunk boundaries at easy finger transitions should lead to slower performance because these finger transitions now have to be used to recall the next chunk (which takes longer than executing the easy finger transitions). We hypothesized that

this would hinder overall performance (Fig. 2C). All participants practiced the same seven sequences (Fig. 2B). Half of the participants in the experimental group were instructed with the aligned chunk structure for the first three sequences and the misaligned chunk structure for the next three sequences (Fig. 2D). For the other half of the participants, the assignment of sequences to aligned and misaligned was reversed. The last sequence (#) was neither misaligned nor aligned under the two chunk structures but was added to ensure that each chunk occurred in at least two different sequences. The counterbalanced design (Fig. 2D) allowed us to draw strong inferences about whether participants' performance was dictated by execution demands (which were identical across participants) or whether it was affected by the chunk structure imposed during the chunk instruction phase (which was different between the two chunk sets). We also included a control group that did not receive any explicit chunk instruction.

Every participant completed 15 training sessions in total (Fig. 2E): one session per day across a 3-wk period. Each session lasted ~1 h, excluding the two initial sessions and the last session that each took 2 h. Participants completed at least 10 blocks of 28 trials per training day. Each block comprised four repetitions of each of the seven sequences.

Days 1–4: chunk instruction and initial sequence learning. For the experimental group, participants were pretrained on one of the two chunk sets on the first day of training (Fig. 2A). Each chunk was associated with a letter of the alphabet (A–K). Participants were explicitly told to learn this association. Each chunk was practiced twice in succession. In the first trial of each pair, the numbers corresponding to the finger presses were shown together with the letter indicating the chunk. In the second trial, the numbers were replaced with stars, such that the participants had to recall the presses solely based on the letter. This trial order was reversed on every second block, such that participants had to first try to recall the sequence with the help of the letter, and then were shown both the letter and the numbers in the second trial. To ensure that participants had memorized the chunks, we added recall blocks at the end of *days 1* and *2*. At the end of the first and the second day, participants were asked to recall and type out the numbers corresponding to the presented letters as quickly and as accurately as possible (letters were randomly presented). At the end of *day 2*, participants could reliably produce the chunks from memory with an average accuracy of 92.7%.
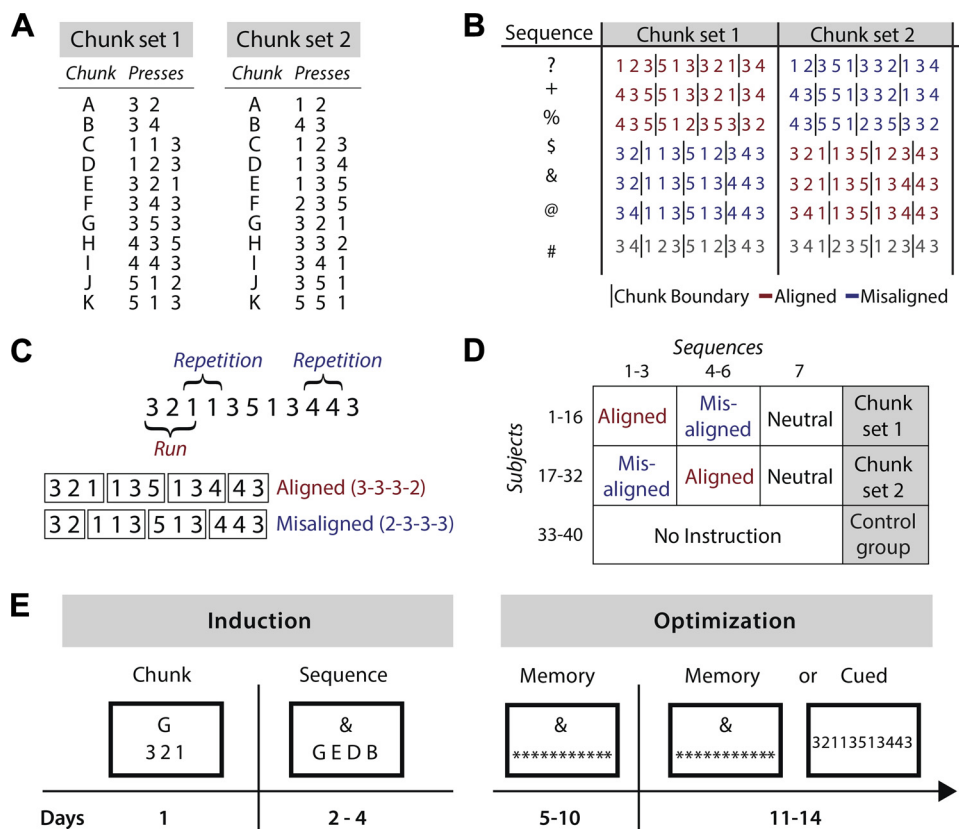
On *day 2*, the experimental participants trained on the seven 11-press sequences. Each sequence was associated with a symbol (e.g., $; Fig. 2B). Each symbol was presented twice in succession, and participants had to perform the sequences from memory using the symbol cue on one trial or with the help of the chunk letters on the next trial. We tested participants' sequence knowledge with a recall block at the end of *days 2–4*. The first two participants did not perform the recall blocks. At the end of *day 4*, participants were able to recall all sequences from memory using the sequence cues with an accuracy of 93.1%.

In contrast, the control group did not receive any chunk training but instead trained directly on the seven 11-press sequences. On *day 1*, they were presented with the 11 digits corresponding to the 11-press sequences. We matched the amount of training across groups by ensuring that all participants were required to produce the same overall number of finger presses. On *day 1*, the control participants were not aware that they would have to memorize the sequences later on. On *days 2–4*, they were instructed to memorize the sequences using the same symbolic sequence cues as the experimental groups, and their memory was tested using recall blocks at the end of each day (*day 4* = 90.2% accuracy). The rest of the experimental design was identical for all groups.

Days 5–10: optimization—memory recall. On *days 5–10*, both experimental and control participants practiced exclusively on the 11-press sequences using the symbolic cues. Chunks were no longer cued. Each sequence cue was presented twice in succession, and participants had to recall the sequence from memory on both trials.

Days 11–14: optimization—memory recall or cued presentation. On the last 4 days of training, half of the experimental participants

Fig. 2. Experimental design. *A*: each participant learned 11 chunks associated with the chunk cues (*A–K*) from one of the chunk sets. *B*: the seven 11-digit sequences that participants trained on. The vertical lines (not shown to the participants) indicate the chunk boundaries induced in training through the chunk set. Sequences were trained with an aligned (red) or misaligned (blue) chunk structure. *C*: example sequence containing a three-digit run and two-digit repetitions. In the aligned structure, the chunk boundaries fell between repetitions, in the misaligned structure the chunk boundary broke up the run. *D*: we counterbalanced across participants which sequences were practiced with which chunk structures. An additional control group was added who did not receive any chunk instruction. *E*: experimental timeline depicting the training at each stage. In the instruction phase, participants memorized chunks and sequences. In the optimization phase, participants trained to perform these sequences as fast as possible from memory. In the last week of training, half of the participants were directly cued with the sequence, whereas the others performed the sequences from memory.

A

Chunk set 1		Chunk set 2	
Chunk	Presses	Chunk	Presses
A	3 2	A	1 2
B	3 4	B	4 3
C	1 1 3	C	1 2 3
D	1 2 3	D	1 3 4
E	3 2 1	E	1 3 5
F	3 4 3	F	2 3 5
G	3 5 3	G	3 2 1
H	4 3 5	H	3 3 2
I	4 4 3	I	3 4 1
J	5 1 2	J	3 5 1
K	5 1 3	K	5 5 1

B

Sequence	Chunk set 1	Chunk set 2
?	1 2 3|5 1 3|3 2 1|3 4	1 2|3 5 1|3 3 2|1 3 4
+	4 3 5|5 1 3|3 2 1|3 4	4 3|5 5 1|3 3 2|1 3 4
%	4 3 5|5 1 2|3 5 3|3 2	4 3|5 5 1|2 3 5|3 3 2
$	3 2|1 1 3|5 1 2|3 4 3	3 2 1|1 3 5|1 2 3|4 3
&	3 2|1 1 3|5 1 3|4 4 3	3 2 1|1 3 5|1 3 4|4 3
@	3 4|1 1 3|5 1 3|4 4 3	3 4 1|1 3 5|1 3 4|4 3
#	3 4|1 2 3|5 1 2|3 4 3	3 4 1|2 3 5|1 2 3|4 3

\| Chunk Boundary   — Aligned   — Misaligned

C

Repetition   *Repetition*

3 2 1 1 3 5 1 3 4 4 3

Run

3 2 1 | 1 3 5 | 1 3 4 | 4 3   Aligned (3-3-3-2)

3 2 | 1 1 3 | 5 1 3 | 4 4 3   Misaligned (2-3-3-3)

D

| | *Sequences* | | | |
	1-3	4-6	7	
1-16	Aligned	Mis-aligned	Neutral	Chunk set 1
17-32	Mis-aligned	Aligned	Neutral	Chunk set 2
33-40	No Instruction			Control group

Subjects

E

Induction		Optimization		
Chunk	Sequence	Memory	Memory	or Cued
G 3 2 1	& G E D B	& ***********	& ***********	3 2 1 1 3 5 1 3 4 4 3

| Days | 1 | 2 - 4 | 5-10 | 11-14 |

---

performed the sequences from memory (as on *days 5–10*), whereas for the other half and for the control participants, we removed the symbolic sequence cue and instead visually presented participants with the complete set of 11 digits that corresponded to the sequences (Fig. 2*E*).

Statistical Analysis

We recorded and analyzed the force measured at each key. For each trial, we calculated movement time (MT, time between the first press and last release) and interpress intervals (IPIs; time between force peaks of two consecutive presses). All analyses were performed using custom-written code in MATLAB (The MathWorks) and the dataframe toolbox (github.com/jdiedrichsen/dataframe). We excluded from our analyses the trials that contained one or more incorrect presses, as well as trials with an MT or a press with an IPI 3 SDs above the mean calculated across all days and participants.

For the correlation analysis in Fig. 5, we split the data for each day, subject, and sequence in half (first vs. second half of each day) and calculated correlations of all possible pairs. We Fisher *z*-transformed the correlations before averaging and performing statistical tests. For plotting the correlations, we then inverse Fisher *z*-transformed the values.

The data were analyzed using mixed-effects analysis of variance (mixed ANOVA), Pearson's correlation, and paired and one-sample *t* tests. All *t* tests were two-sided unless specified otherwise. A probability threshold of $P < 0.05$ for the rejection of the null hypothesis was used for all statistical tests.

RESULTS

Over 15 days we trained 32 participants to produce sequences of 11 isometric keypresses from memory on a keyboard-like device. Participants were rewarded with points for executing sequences as fast as possible while keeping the proportion of incorrect keypresses in each block of trials below 15%. We maintained the participants' motivation by gradually decreasing the movement time (MT) threshold at which they received points.

We manipulated how participants memorized the sequences by splitting the sequences into several chunks, each composed of two to three keypresses. The aim was to test whether the different ways of chunking (hereafter "chunk structures") imposed through the chunk training in the instruction phase (METHODS; Fig. 2*B*) would affect performance optimization in the subsequent 2 wk of training. Each sequence could be chunked in an aligned or misaligned fashion, predicted to lead to beneficial or detrimental performance, respectively (METHODS; Fig. 2*C*). All participants practiced the same seven sequences but differed in the chunking instructions they received for each sequence.

Chunk Instruction Induces a Stable Movement Pattern

To assess whether the imposed chunk structures influenced participants' motor behavior, we examined interpress time intervals (IPIs). An increased IPI is commonly taken as a sign of a chunk boundary, as the cognitive processes (memory recall, action selection) involved in switching from one chunk to another require additional time (Verwey 1999; Verwey et al. 2010). Hence, we would expect our participants to exhibit shorter IPIs between keypresses that belonged to a chunk imposed during *day 1* (within-chunk IPIs) and larger IPIs for the boundaries between chunks (between-chunks IPIs). For this analysis, we pooled the data from all sequences irrespective of instruction (misaligned vs. aligned). We indeed found significantly longer between-chunks IPIs compared with within-chunk
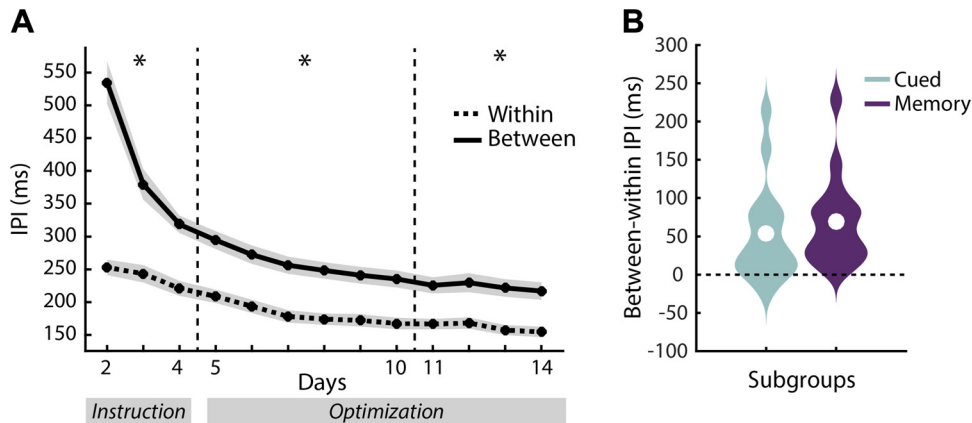
A

B

IPIs in the first few days of training [Fig. 3*A*: *days 2–4*: t $(31) = 7.728$, $P = 5.098\ e−09$], suggesting that our manipulation was successful in inducing a temporally specific pattern of keypresses.

In the optimization phase, we ceased to cue sequences using the alphabetic letters associated with the chunks. Instead, partic-ipants were asked to recall the entire 11-keypress sequences from memory in response to symbolic sequence cues (e.g., "$"). Across *days 5–10*, the within-chunk and between-chunks IPIs were still significantly different from each other [$t(31) = 7.165$, $P = 2.351\ e−08$; Fig. 3*A*]. This difference cannot be attributed to differences in performance difficulty of the finger transitions, as the within-chunk IPIs for one half of the participants were the between-chunks IPIs for the other half and vice versa (Fig. 2*B*). IPIs that were within-chunk for all participants (e.g., the first and last IPI of a sequence) were excluded from this analysis.

In the last 4 days of training, we tested whether the slower IPIs at chunk boundaries were because of the fact that the sequences needed to be recalled from memory. Half of the par-ticipants continued to perform the sequences from memory, whereas the other half were cued using the numbers that indi-cated the necessary keypresses (Fig. 2*E*), therefore removing any memory recall demands. Both the memory [$t(15) = 4.865$, $P = 2.059\ e−04$; Fig. 3*B*] and the cued subgroup [$t(15) = 3.403$, $P = 0.004$] showed a significant difference between the within-chunk and between-chunks IPIs. There was no reliable differ-ence between the two subgroups in this effect [$t(30) = −0.749$, $P = 0.460$]. Thus, removing the requirement for memory recall did not abolish chunking. Because none of the subsequent anal-yses showed any significant difference between the two sub-groups, we will report their combined results for the remainder of the article.

Misaligned Chunk Structure Impairs Performance

We then asked whether the two types of chunk instructions that were used for each sequence (counterbalanced across par-ticipants) would lead to measurable differences in performance. We designed chunk structures that were either aligned or mis-aligned with the basic execution-level constraints (see METHODS) and predicted that these structures would either improve or impede performance. Each experimental participant learned three of the seven sequences with a misaligned chunk struc-ture and three sequences with an aligned chunk structure, with the assignment counterbalanced across participants (Fig.

2*D*). Therefore, all participants practiced the same seven sequences but differed in which chunk instructions they received. This counterbalanced design allowed us to compare execution speed between aligned and misaligned sequences within each participant.

To test our prediction that training with the misaligned chunk structure would lead to poorer performance, we meas-ured participants' movement time (MT) by estimating the time between the first finger press and the last finger release. For each participant, we then calculated the difference in average speed between the aligned and the misaligned instructed sequences. As predicted, in the instruction phase, sequences instructed with the misaligned chunk structure were performed slower than the sequences instructed with the aligned chunk structure (Fig. 4*A*) [one-sample t test: t $(31) = 2.693$, $P = 0.006$]. Hence, we were able to manipulate not only how participants performed a sequence but also how well they could perform it.

Next, we wanted to examine what factors influenced the dif-ference in speed we observed. To determine how beneficial it was to have a finger run (3 adjacent presses in either descending or ascending order, e.g., 123) preserved within a chunk, rather than separated by a chunk boundary, we selected all IPIs that could be either within or between a chunk (excluding the IPIs that were within-chunk for both aligned and misaligned struc-tures). For the within-chunk IPIs, we compared the average IPI for transitions that occurred in a run (e.g., between 1 & 2 and 2 & 3) with the average IPI for transitions outside of a run (e.g., 2 & 4 or 5 & 1). This gives us an idea of how beneficial it is for a finger run to be present within a chunk compared with when there is no run present. We then compared this measurement to how advantageous a run is when it is separated by a chunk boundary. In other words, we compared the advantage of having a finger run within a chunk to having a finger run that is distrib-uted across chunks. We found a significant advantage of 28.6 ms [SD = 44.6; one-sided t test: $t(31) = 3.624$, $P = 5.137\ e−4$]. Similarly, we calculated the cost of a finger repetition within a chunk, as compared with the cost of a finger repetition between chunks, and found an average difference of 16 ms (SD = 68.1), a nonsignificant difference [one-sided t test: $t(31) = 1.331$, $P = 0.097$]. An additional factor that influenced participants' speed was whether the two-digit chunk was placed in the beginning (mis-aligned) or the end (aligned) of the sequence. We evaluated this factor by averaging the second and second-to-last IPI in each
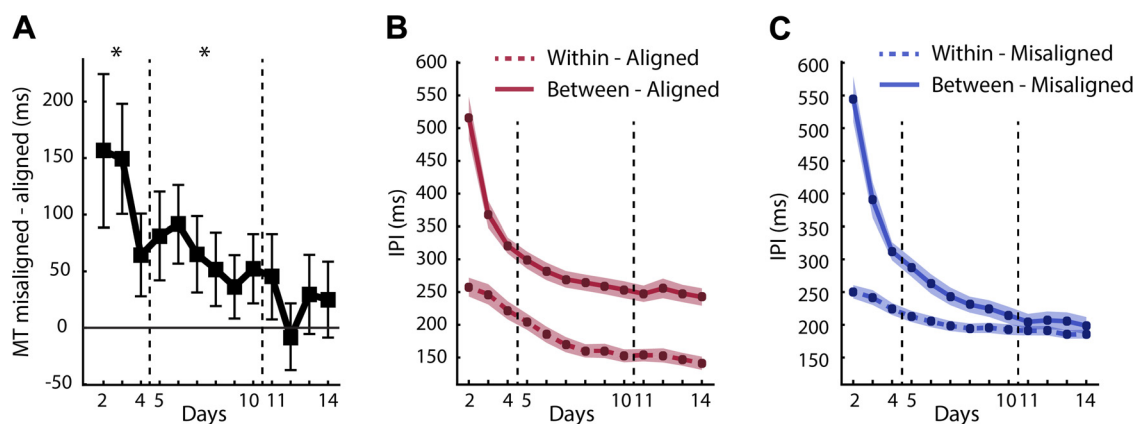
Fig. 4. Change in chunk structure and performance for aligned and misaligned instructed sequences. *A*: differences in movement time (MT) between sequences instructed with an aligned or misaligned chunk structure. *Significant difference from 0 (no difference). *B*: within-chunk or between-chunks interpress intervals (IPIs) across training days for the sequences instructed with the aligned chunk structure. *C*: within-chunk or between-chunks IPIs across training days for the sequences instructed with the misaligned chunk structure. Error bars denote between-subjects standard error.

sequence, as one of them was within-chunk and one was between-chunks for each sequence. This comparison showed a significant advantage of 24.7 ms (SD = 60.0) for the aligned chunk structure [one-sided *t* test: $t(31) = 2.330$, $P = 0.013$]. These results suggest that multiple factors led to an MT advantage for sequences that were instructed with an aligned versus misaligned chunk structure.

The difference in MT we found in the first week was maintained in the second week of training [*days 5–10*: $t(31) = 2.313$, $P = 0.014$]. However, this speed difference was no longer statistically reliable in the last 4 days of training [*days 11–14*: $t(31) = 0.764$, $P = 0.225$]. This suggests that participants ultimately were able to overcome the performance detriment that we imposed through the initial chunk instructions.

To determine whether receiving the aligned chunk instruction was more beneficial to performance than not receiving a chunk instruction at all, we tested an additional control group ($N = 8$). This group did not have to explicitly learn chunks but rather trained on the entire sequences from the beginning (see METHODS for details). We compared the performance of this group with that of the experimental group for *days 5–10*, during which all groups had to perform the sequences from memory. During these days, the control group performed, on average, 63.5 ms [standard error (SE) = 223.2 ms] slower than the experimental group on the aligned sequences, and nearly identical (0.5 ms slower, SE = 215.3 ms) compared with the performance of the experimental group on the misaligned sequences. However, neither of these contrasts reached statistical significance [aligned vs. control: $t(38) = -0.285$, $P = 0.778$; misaligned vs. control: $t(38) = -0.002$, $P = 0.998$].

Misaligned Chunk Structure Is Changed More Rapidly

To investigate how participants overcame the detrimental influence of the misaligned chunk structure, we separated the IPI analysis (Fig. 3*A*) by whether the intervals came from sequences that were instructed using an aligned or misaligned structure. The difference between within-chunk and between-chunks IPIs for sequences instructed with the aligned chunk structure was stable over the entire training period (Fig. 4*B*). In contrast, for the misaligned structure, the difference between the within-chunk and between-chunks IPIs started to disappear late in learning (Fig. 4*C*). The three-way day × within/

between × aligned/misaligned interaction was significant [$F(12,372) = 19.790$, $P = 1\ e{-}16$]. Thus, participants diverged from the misaligned chunk structure while maintaining the aligned chunk structure.

To understand these changes in more detail, we investigated the entire pattern of IPIs produced by the participants for each sequence. In a first analysis, we correlated the participants' IPI pattern of each day to the pattern produced on *day 2* (Fig. 5*A*, see METHODS for details). This analysis shows how far participants diverged from their initial chunking pattern with training. The comparison between the aligned and misaligned instructed sequences confirmed our previous observation that participants diverged more from the misaligned instruction [day × instruction: $F(11,330) = 4.348 = P = 4.352\ e{-}06$]. The analysis also demonstrates that the control group significantly diverged from their second-day IPI pattern with training [day: $F(11,77) = 30.209$, $P < 0.0001$].

Importantly, our data show that this drift was not because of participants becoming more variable in their performance. To investigate the stability of the temporal structure within each day, participant, and sequence, we correlated the average IPI patterns across the first half and second half of each day. To test for a systematic change of stability across training, we fitted a linear regression separately to each participant's correlation results and compared the resulting slope values to zero. We found that within-subject correlations increased over the course of training for the aligned instructed sequences [$t(31) = 4.204$, $P = 3.071\ e{-}05$] as well as for the control group [$t(31) = 2.874$, $P = 0.024$; Fig. 5*B*]. For the misaligned instructed sequences, the increase failed to reach significance [$t(31) = 1.9744$, $P = 0.0573$; Fig. 5*B*]. We also found that the chunking pattern for the misaligned instructed sequences was less stable than for the aligned instructed sequences [$t(31) = 2.952$, $P = 0.006$]. Overall, however, correlations were very high ($r > 0.9$), showing that participants adopted a relatively stable temporal structure for each sequence.

Given that participants converged on a stable IPI pattern for each sequence, we asked whether this pattern was the same across participants, or whether individuals found idiosyncratic solutions. To explore this question, we again used the average IPI pattern for each half day, but now correlated these patterns
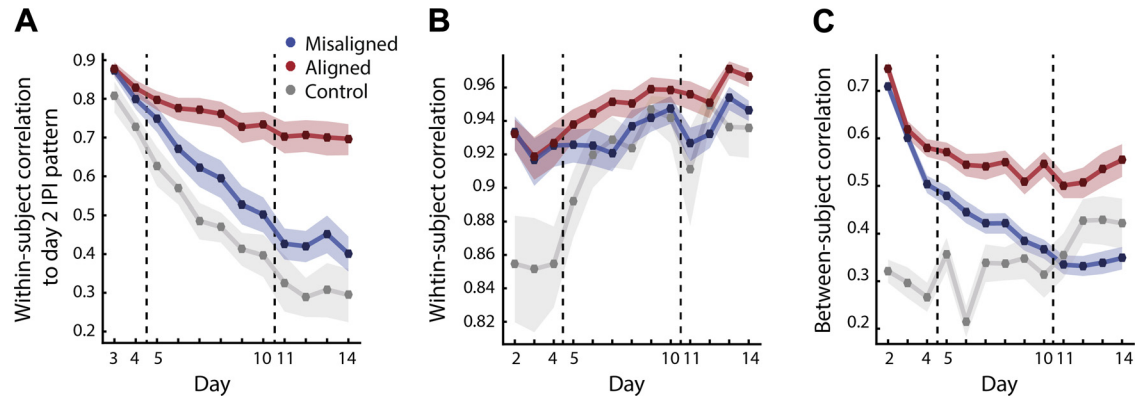
Fig. 5. Changes of interpress interval (IPI) pattern across training. *A*: within-subject correlation of IPI pattern of *day 2* with the IPI pattern of each subsequent day, averaged across the first and second half of blocks. Correlation separated for misaligned and aligned instructed sequences and control group. *B*: within-subject IPI pattern per day between first and second half of data. *C*: between-subjects IPI pattern separated by instruction across days.

with those of any of the other participants. This approach enabled us to directly compare how similar two participants performed the same sequence in a session (between-subjects correlation) with how consistent a single participant performed that same sequence (within-subject correlation). For the experimental group, we found that the between-subjects correlation (Fig. 5*C*) was substantially lower than the within-subject correlation [$t(31) = 19.664$, $P < 0.0001$] at the end of training (*day 14*). This suggests that participants adopted chunk structures at the end of training that were stable but quite different across participants. This was especially true for the misaligned instructed sequences, which showed a lower between-subjects correlation than the aligned instructed sequences on the last day of training [$t(31) = -8.211$, $P = 2.834\,e-09$; Fig. 5*B*]. Similarly, the control group also shows much higher within-subject than between-subjects correlation of the IPI patterns [$t(7) = -19.119$, $P = 2.666\,e-07$]. Together, these results show that participants, independent of chunk instruction, changed their IPI patterns systematically over training, converging on idiosyncratic, but individually stable, temporal patterns of performance.

DISCUSSION

In this study, we used chunking as a tool to investigate the role of instructions on skill learning. We influenced the structure of the initial declarative sequence representation by manipulating how participants memorized them (Park et al. 2004). Moreover, by experimentally imposing two different chunk structures on the same physical sequence, one that was beneficial and one that was detrimental to performance, we could make causal inferences about the effects of chunking on motor skill development. This is an important advance over previous observational studies (Ramkumar et al. 2016; Wright et al. 2010; Wymbs et al. 2012), which did not experimentally control how participants chose to chunk the sequences.

We report three main results. First, consistent with previous studies (de Kleine and Verwey 2009; Verwey and Dronkert 1996; Verwey et al. 2009, 2010), our data demonstrate that a stable chunking pattern can be induced through cognitive manipulations during the initial stages of sequence learning. Importantly, participants did not completely overcome this imposed chunk structure, and the chunking structure remained stable, even when the task changed from a memory-guided to a

stimulus-guided task. Thus, the initial chunk instructions led to the formation of specific movement patterns.

Second, we were able to induce chunking patterns that differentially affect participants' performance. To do so, we designed two different ways of instructing the sequence, one aligned and the other misaligned with execution-level constraints that were identified using a separate participant population. Using this manipulation, we were able to induce a performance difference in the beginning of practice, which was still observed during the second week of practice but disappeared in the last week. Although these results clearly show that instructions can systematically impact performance, the comparison with participants who trained without explicit chunking instructions does not allow firm conclusions on whether this difference was caused by the aligned instruction facilitating performance, the misaligned instruction impeding performance, or both.

Finally, more detailed analyses of the interpress interval patterns showed that participants followed the beneficial chunk instructions throughout the entire training period but changed their chunking pattern for the misaligned instructed sequences. We also demonstrate that participants did not all converge on the same chunking pattern after abandoning the misaligned instructions but rather found an idiosyncratic chunking structure for each sequence. These solutions differed across participants but were relatively stable within each participant at the end of training. Similar observations were made for the control group. The stabilization of IPI patterns that we observed over the course of training can be compared with the development of an invariant temporal and spectral structure in birdsong, a process that has been termed "crystallization" (Brainard and Doupe 2002).

An alternative interpretation of the data is that with training, participants' temporal interpress interval patterns are primarily driven by execution-level constraints of the sequences rather than by chunking. This would mean that chunking is abolished with training and the remaining regularities we observe are due to physical constraints (i.e., how fast can each finger transition be executed). Some aspects of our findings, however, speak against this possibility. Execution-level constraints have a relatively high ($r = 0.689$) intersubject correlation (see *Baseline Study for Measuring Execution-Level Constraints*), and we would have expected a similarly high correlation for the IPI

patterns. Our finding that the between-subjects IPI correlation is substantially lower and also differed across chunk instructions therefore speaks against this possibility. This line of reasoning would also suggest that participants who did not receive any explicit chunk instructions (the control group) developed stable chunking patterns with training, supporting claims by previous studies (Ramkumar et al. 2016; Wright et al. 2010; Wymbs et al. 2012).

The characteristics of the stable motor patterns we observed make them similar to "habits." Habits are defined as a highly entrenched behavioral pattern that resists change through retraining (Ashby et al. 2003; Dezfouli and Balleine 2012; Dolan and Dayan 2013; Graybiel 2008; Graybiel and Grafton 2015; Hardwick et al. 2019; Hélie et al. 2010; Jager 2003; Robbins and Costa 2017; Seger and Spiering 2011; Smith and Graybiel 2013a), even if they have become maladaptive. Most articles on habits (Jog et al. 1999; Robbins and Costa 2017; Smith and Graybiel 2014, 2016; Wickens et al. 2007) have focused on habits in the context of action selection, i.e., choosing what action to perform. In contrast, our experiment addresses the question of habits in motor performance, i.e., habits that influence how to perform a chosen action. For example, a tennis player could be influenced by a habitual pattern in action selection, whereby she always chooses a forehand over a backhand to return a serve. At the same time, she could be influenced by a motor habit, whereby she executes the forehand without rotating her hips. In support of this idea, we showed that we could induce a stable performance pattern that can be observed even after weeks of training, that these performance patterns crystallized over the course of training, and that changes in task demands did not lead to behavioral modifications. Therefore, we believe that studying chunking can provide valuable insights into the neural systems underlying motor habits. Indeed, it has recently been suggested that chunking plays an integral role in the formation and expression of habits (Dezfouli et al. 2014; Graybiel 2008) and is neurally represented in the dorsal lateral striatum as action "start and stop signals" (Barnes et al. 2005; Graybiel 1998; Jin et al. 2014; Smith and Graybiel 2013a, 2014).

The establishment of a paradigm that allows us to not only cognitively influence participant movement patterns but also influence their behavioral impact on performance will enable us to explore ways to encourage learners to change their current movement pattern, especially if it is disadvantageous. Although our deliberate attempt at modifying participants' behavior by changing the task from a memory-based to a stimulus-based task was ultimately not successful, there are many other techniques that could be used. In many disciplines, teachers have developed ways to help students overcome habits. For instance, the Hanon piano exercise helps students play difficult passages of a musical piece by breaking up learned phrases into new chunks to explore different rhythms. Playing a passage slower than intended has also been suggested to break habits (Chang 2016). These techniques suggest that changes in context can be helpful in breaking motor habits after they have formed.

In conclusion, we were able to use a sequence chunking paradigm to impose specific movement patterns on participants' behavior that were stable across 3 wk of training. Furthermore, by aligning the imposed chunking patterns with beneficial or detrimental finger transitions, we were able to impact participants'

performance speed. Interestingly, although participants maintained the beneficial chunking pattern throughout the entire training period, they were able to abandon the detrimental pattern to overcome the imposed performance detriment.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

N.J.P., A.Y., and J.D. conceived and designed research; N.J.P. performed experiments; N.J.P., and J.D. analyzed data; N.J.P., P.L.G., and J.D. interpreted results of experiments; N.J.P. prepared figures; N.J.P. drafted manuscript; N.J.P., P.L.G., and J.D. edited and revised manuscript; N.J.P., A.Y., P.L.G., and J.D. approved final version of manuscript.

REFERENCES

Abrahamse EL, **Ruitenberg MFL**, **de Kleine E**, **Verwey WB.** Control of automated behavior: insights from the discrete sequence production task. *Front Hum Neurosci* 7: 82, 2013. doi:10.3389/fnhum.2013.00082.

Ashby FG, **Ell SW**, **Waldron EM.** Procedural learning in perceptual categorization. *Mem Cognit* 31: 1114–1125, 2003. doi:10.3758/BF03196132.

Barnes TD, **Kubota Y**, **Hu D**, **Jin DZ**, **Graybiel AM.** Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* 437: 1158–1161, 2005. doi:10.1038/nature04053.

Brainard MS, **Doupe AJ.** What songbirds teach us about learning. *Nature* 417: 351–358, 2002. doi:10.1038/417351a.

Chang CC. *Fundamentals of Piano Practice* (3rd ed.). Tampa, FL: CreateSpace Independent Publishing Platform, 2016.

Dezfouli A, **Balleine BW.** Habits, action sequences and reinforcement learning. *Eur J Neurosci* 35: 1036–1051, 2012. doi:10.1111/j.1460-9568.2012.08050.x.

Dezfouli A, **Lingawi NW**, **Balleine BW.** Habits as action sequences: hierarchical action control and changes in outcome value. *Philos Trans R Soc Lond B Biol Sci* 369: 20130482, 2014. doi:10.1098/rstb.2013.0482.

Diedrichsen J, **Kornysheva K.** Motor skill learning between selection and execution. *Trends Cogn Sci* 19: 227–233, 2015. doi:10.1016/j.tics.2015.02.003.

Dolan RJ, **Dayan P.** Goals and habits in the brain. *Neuron* 80: 312–325, 2013. doi:10.1016/j.neuron.2013.09.007.

Graybiel AM. The basal ganglia and chunking of action repertoires. *Neurobiol Learn Mem* 70: 119–136, 1998. doi:10.1006/nlme.1998.3843.

Graybiel AM. Habits, rituals, and the evaluative brain. *Annu Rev Neurosci* 31: 359–387, 2008. doi:10.1146/annurev.neuro.29.051605.112851.

Graybiel AM, **Grafton ST.** The striatum: where skills and habits meet. *Cold Spring Harb Perspect Biol* 7: a021691, 2015. doi:10.1101/cshperspect. a021691.

Green TD, **Flowers JH.** Implicit versus explicit learning processes in a probabilistic, continuous fine-motor catching task. *J Mot Behav* 23: 293–300, 1991. doi:10.1080/00222895.1991.9942040.

Halford GS, **Wilson WH**, **Phillips S.** Processing capacity defined by relational complexity: implications for comparative, developmental, and cognitive psychology. *Behav Brain Sci* 21: 803–831, 1998. doi:10.1017/S0140525X98001769.

Hardwick RM, **Forrence AD**, **Krakauer JW**, **Haith AM.** Time-dependent competition between goal-directed and habitual response preparation. *Nat Hum Behav* 3: 1252–1262, 2019. doi:10.1038/s41562-019-0725-0.

Hélie S, **Waldschmidt JG**, **Ashby FG.** Automaticity in rule-based and information-integration categorization. *Atten Percept Psychophys* 72: 1013–1031, 2010. doi:10.3758/APP.72.4.1013.

Hodges NJ, **Franks IM.** Modelling coaching practice: the role of instruction and demonstration. *J Sports Sci* 20: 793–811, 2002. doi:10.1080/026404102320675648.

Jager W. Breaking 'bad habits': a dynamical perspective on habit. In: *Human Decision Making and Environmental Perception. Understanding and Assisting Human Decision Making in Real-Life Settings*, edited by Hendrickx L, Jager W, Steg L. Groningen, Netherlands: Liber Amicorum for Charles Vlek, 2003, p. 149–160.

Jin X, **Tecuapetla F**, **Costa RM.** Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nat Neurosci* 17: 423–430, 2014. doi:10.1038/nn.3632.

Jog MS, **Kubota Y**, **Connolly CI**, **Hillegaart V**, **Graybiel AM.** Building neural representations of habits. *Science* 286: 1745–1749, 1999. doi:10.1126/science.286.5445.1745.

de Kleine E, **Verwey WB.** Representations underlying skill in the discrete sequence production task: effect of hand used and hand position. *Psychol Res* 73: 685–694, 2009. doi:10.1007/s00426-008-0174-2.

Masters RSW. Knowledge, knerves and know-how: the role of explicit versus implicit knowledge in the breakdown of a complex motor skill under pressure. *Br J Psychol* 83: 343–358, 1992. doi:10.1111/j.2044-8295.1992.tb02446.x.

Meier C, **Frank C**, **Gröben B**, **Schack T.** Verbal instructions and motor learning: how analogy and explicit instructions influence the development of mental representations and tennis serve performance. *Front Psychol* 11: 2, 2020. doi:10.3389/fpsyg.2020.00002.

Miller GA. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 63: 81–97, 1956. doi:10.1037/h0043158.

Park J-H, **Wilde H**, **Shea CH.** Part-whole practice of movement sequences. *J Mot Behav* 36: 51–61, 2004. doi:10.3200/JMBR.36.1.51-61.

Ramkumar P, **Acuna DE**, **Berniker M**, **Grafton ST**, **Turner RS**, **Kording KP.** Chunking as the result of an efficiency computation trade-off. *Nat Commun* 7: 12176, 2016. doi:10.1038/ncomms12176.

Robbins TW, **Costa RM.** Habits. *Curr Biol* 27: R1200–R1206, 2017. doi:10.1016/j.cub.2017.09.060.

Seger CA, **Spiering BJ.** A critical review of habit learning and the basal ganglia. *Front Syst Neurosci* 5: 66, 2011. doi:10.3389/fnsys.2011.00066.

Smith KS, **Graybiel AM.** Using optogenetics to study habits. *Brain Res* 1511: 102–114, 2013a. doi:10.1016/j.brainres.2013.01.008.

Smith KS, **Graybiel AM.** Investigating habits: strategies, technologies and models. *Front Behav Neurosci* 8: 39, 2014. doi:10.3389/fnbeh.2014.00039.

Smith KS, **Graybiel AM.** Habit formation coincides with shifts in reinforcement representations in the sensorimotor striatum. *J Neurophysiol* 115: 1487–1498, 2016. doi:10.1152/jn.00925.2015.

Solopchuk O, **Alamia A**, **Olivier E**, **Zénon A.** Chunking improves symbolic sequence processing and relies on working memory gating mechanisms. *Learn Mem* 23: 108–112, 2016. doi:10.1101/lm.041277.115.

Verwey WB. Buffer loading and chunking in sequential keypressing. *J Exp Psychol Hum Percept Perform* 22: 544–562, 1996. doi:10.1037/0096-1523.22.3.544.

Verwey WB. Evidence for a multistage model of practice in a sequential movement task. *J Exp Psychol Hum Percept Perform* 25: 1693–1708, 1999. doi:10.1037/0096-1523.25.6.1693.

Verwey WB. Concatenating familiar movement sequences: the versatile cognitive processor. *Acta Psychol (Amst)* 106: 69–95, 2001. doi:10.1016/S0001-6918(00)00027-5.

Verwey WB, **Abrahamse EL**, **Jiménez L.** Segmentation of short keying sequences does not spontaneously transfer to other sequences. *Hum Mov Sci* 28: 348–361, 2009. doi:10.1016/j.humov.2008.10.004.

Verwey WB, **Abrahamse EL**, **de Kleine E.** Cognitive processing in new and practiced discrete keying sequences. *Front Psychol* 1: 32, 2010. doi:10.3389/fpsyg.2010.00032.

Verwey WB, **Dronkert Y.** Practicing a structured continuous key-pressing task: motor chunking or rhythm consolidation? *J Mot Behav* 28: 71–79, 1996. doi:10.1080/00222895.1996.9941735.

Verwey WB, **Eikelboom T.** Evidence for lasting sequence segmentation in the discrete sequence-production task. *J Mot Behav* 35: 171–181, 2003. doi:10.1080/00222890309602131.

Wickens JR, **Horvitz JC**, **Costa RM**, **Killcross S.** Dopaminergic mechanisms in actions and habits. *J Neurosci* 27: 8181–8183, 2007. doi:10.1523/JNEUROSCI.1671-07.2007.

Wiestler T, **Diedrichsen J.** Skill learning strengthens cortical representations of motor sequences. *eLife* 2: e00801, 2013. doi:10.7554/eLife.00801.

Wiestler T, **Waters-Metenier S**, **Diedrichsen J.** Effector-independent motor sequence representations exist in extrinsic and intrinsic reference frames. *J Neurosci* 34: 5054–5064, 2014. doi:10.1523/JNEUROSCI.5363-13.2014.

Wong AL, **Lindquist MA**, **Haith AM**, **Krakauer JW.** Explicit knowledge enhances motor vigor and performance: motivation versus practice in sequence tasks. *J Neurophysiol* 114: 219–232, 2015. doi:10.1152/jn.00218.2015.

Wright DL, **Rhee J-H**, **Vaculin A.** Offline improvement during motor sequence learning is not restricted to developing motor chunks. *J Mot Behav* 42: 317–324, 2010. doi:10.1080/00222895.2010.510543.

Wulf G, **Höß M**, **Prinz W.** Instructions for motor learning: differential effects of internal versus external focus of attention. *J Mot Behav* 30: 169–179, 1998. doi:10.1080/00222899809601334.

Wymbs NF, **Bassett DS**, **Mucha PJ**, **Porter MA**, **Grafton ST.** Differential recruitment of the sensorimotor putamen and frontoparietal cortex during motor chunking in humans. *Neuron* 74: 936–946, 2012. doi:10.1016/j.neuron.2012.03.038.

Yokoi A, **Bai W**, **Diedrichsen J.** Restricted transfer of learning between unimanual and bimanual finger sequences. *J Neurophysiol* 117: 1043–1051, 2017. doi:10.1152/jn.00387.2016.