Buffer Loading and Chunking in Sequential Keypressing

Willem B. Verwey
TNO Human Factors Research Institute

Thirty-six participants practiced a task in which they continuously cycled through a fixed series of nine keypresses, each carried out by a single finger (cf. Keele & Summers, 1976). The results of the first experimental phase, the practice phase, support the notion that pauses between successive keypresses at fixed locations induces the development of integrated sequence representations (i.e., motor chunks) and reject the idea that a rhythm is learned. When different sequences were produced in the transfer phase, performance dropped considerably unless the sequence was relatively short and there was ample time for preparation. This demonstrates that motor chunks are content specific and that the absence of motor chunks shows when there is no time for advance loading of the motor buffer or the capacity of the motor buffer is insufficient to contain the entire keypressing sequence.

Since the early days of research on human motor behavior it has been known that practice is the predominant factor in shaping performance. Yet, the basic mechanisms underlying practice effects are still largely unknown. One classic but still influential notion is that, with practice, elements pertaining to a specific task are encapsulated in integrated chunks which, then, can be handled as a single and more simple representation (Miller, 1956). As regards the execution of movement sequences, this could mean that chunks are constructed by combining representations of elementary, sometimes innate, movements or movement patterns (e.g., Adams, 1984; Book 1908; Bruner, 1973; Fentress, 1984; Keele, 1986; Lashley, 1951; Miller, Galanter, & Pribram, 1960; Paillard, 1960). The benefit of motor chunks would lie in the associated reduction of storage and retrieval capacity (see, e.g., Gallistel, 1980; Jones, 1981; Newell & Rosenbloom, 1981).

Relatively little is known about the specificity of practice on the production of movement sequences. The main purpose of this article is to investigate the notion of motor chunking and pursue its relation to the production of relatively short movement sequences and motor storage. To that end, the first part of this article describes a practice phase that investigated the development of motor chunks in a sequential keypressing task in which the inclusion of a few long response stimulus intervals (RSIs) at fixed positions was assumed to determine the boundaries between developing chunks. In the second part of this article, the transfer phase addresses the specificity of practice as implied in the chunking notion.

Verwey and Dronkert (1996) investigated chunk development in a task involving prolonged reproduction of the same sequence of keypresses. Each trial consisted of nine keypressing responses to nine stimuli that were carried out in rapid succession with nine fingers. Each trial was immediately followed by the next identical trial. In this way the task consisted of continuous cycling through the same nine keypresses. In the heavily practiced *Structured* condition, each trial had two or three RSIs that partitioned a sequence into two response groups for half of the participants (45 condition) and into three response groups for the remaining participants (333 condition; the first RSI separated successive trials). In addition, participants occasionally carried out a few probe trials in the *Unstructured* condition that had no RSIs at all. Thus, a new stimulus immediately followed the depression of the preceding key. This Unstructured condition was used to test for chunk development, which would be indicated by the occurrence of relatively long response times at the locations where the long RSIs had occurred in the Structured condition. The results showed that Unstructured response times,[1] clearly and increasingly, reflected the RSIs in the Structured condition. In other words, responses in the Unstructured condition were slower at the positions of the RSIs in the Structured condition. This accords with the notion that successive motor chunks are separated by relatively long intervals (Gee & Grosjean, 1983; Machlis, 1977; Rosenbaum, Kenny, & Derr, 1983; Sternberg, Knoll, & Turock, 1990). According to Verwey and Dronkert (1996) these motor chunks had developed as a consequence of repeatedly and consistently loading the motor buffer (Henry & Rogers, 1960; Sternberg, Monsell, Knoll, & Wright, 1978) because of the formation of associations between ensuing responses in a response group (Brown & Carr, 1989; Fischman & Lim, 1991; MacKay, 1982, 1987). The alternative explanation that grouping was caused by the learning of interkey timing (i.e., a fixed rhythm; Keele &

---

[1] Given the fact that the same sequence was repeated over and over again, this article refers to the more neutral term *response time* rather than the more common *reaction time* because, with practice, responses could be anticipated rather than given in reaction to the stimuli.

Summers, 1976) was rejected because this model could neither predict the pattern of results in the 45 condition nor could it explain why ratios between group-start and within-group intervals exceeded 2:1. The 2:1 ratio would result from using a fixed time pattern in which basically all elements in the output occur at equal time intervals between successive elements. To represent the sequence as a binary tree, which is assumed to be commonly used for representing hierarchically structured sequences, "blank" elements (i.e., covert sequence elements) are inserted at appropriate points. In this task, these points would be at the locations of the RSIs in the Structured condition.

This study aimed at replicating and extending the results obtained by Verwey and Dronkert (1996). The reasoning in this article rests on the following two notions:

1. *Buffer loading:* If there is sufficient time available, sequences can be programed in advance in a short-term motor buffer (Henry & Rogers, 1960; Sternberg et al., 1978). In principle, buffer loading is not specific to the content of a certain sequence. Longer sequences require more programming time than shorter sequences (i.e., the complexity effect) and the average within-group response times increase with sequence length.

2. *Chunking:* Motor chunks develop as a result of repeatedly filling the motor buffer with the same elements through the gradual development of interelement associations (Brown & Carr, 1989; MacKay, 1982; Verwey & Dronkert, 1996). Hence, chunks are specific with regard to the sequence they represent. A necessary assumption here is that loading a buffer with a chunk requires less time than when, in the absence of a chunk, the individual sequence elements in the motor buffer need to be selected and loaded one by one. Once loaded, both types of sequences can be executed rapidly. Thus, execution rate of short prepared sequences may be relatively unaffected by the amount of practice. This article addresses the relative contributions of buffer loading and chunking in Verwey and Dronkert's (1996) continuous keypressing task.

Practice Phase

The buffer loading and chunking notions allow a prediction for Verwey and Dronkert's (1996) keypressing task that has not yet been tested: When performance relies on advance programming (i.e., buffer loading), performance will improve rapidly, whereas improvement will be slower when there is no opportunity for advance preparation because, in that case, performance relies on the existence of chunks. The point is that general mechanisms such as buffer loading are easily learned and are probably available to every adult, whereas chunks are content specific and develop only gradually (MacKay, 1982; Verwey, 1994). Hence, in the Structured condition the possibility of preparing a sequence in the motor buffer renders the existence of a motor chunk relatively unimportant for performance. In the Unstructured condition, there is no opportunity for advance preparation—due to the absence of RSIs—so that performance will rely more on the existence of chunks. Because chunks are as-

sumed to develop gradually, responses at the start of a response group and those within a response group should slowly reach their asymptote in the Unstructured condition. Verwey and Dronkert (1996) observed that, after practice, Unstructured response times were generally longer than Structured response times. From the notion that chunks play a dominant role in the production of Unstructured response groups, there seems to be no theoretical reason for this difference other than limited practice. Thus, after extended practice, Unstructured response times should approach Structured response times again.

With respect to the capacity of the motor buffer, note that various authors suggested that practice has the effect that longer sequences can be programmed in the motor buffer (Hulstijn & Van Galen, 1983, 1988; Teulings, Mullins, & Stelmach, 1986). One might argue that this is possible because chunks load the buffer less than ad hoc generated sequence representations. If so, evidence for the development of chunks need not only occur in the Unstructured condition: Chunk development could also be indicated by disappearance of the long within-group response times in the Structured condition which would indicate buffer reloading. Indeed, Verwey (1994) found that the third response in a four-key sequence was relatively slow with little practice but that this disadvantage disappeared with practice. Schneider and Fisk (1983) and Verwey and Dronkert (1996) reported a similar effect of practice.

A different but, in this type of task, closely related issue concerns the development of content-aspecific effect of practice, *concurrent preparation.*[2] In the Verwey and Dronkert (1996) study, Unstructured within-group response times were longer than Structured within-group response times. This was attributed to the fact that chunks had not yet fully developed and responses in the Unstructured condition still relied to some extent on one-by-one execution. However, detailed analysis of the Verwey and Dronkert data showed that the slowing of within-group response times in the Unstructured condition was more pronounced in the 333 than in the 45 condition. Why would longer response groups suffer less from the absence of RSIs than shorter response groups? One explanation rests on the notions that (a) the time to initiate a chunk is independent of its size (as suggested by the gradual reduction of the complexity effect with practice; Fischman & Lim, 1991; Hulstijn & Van Galen, 1983; Teulings et al., 1986; Verwey, 1994; Wing, 1978), and that (b) longer response groups are less hampered by concurrent preparation than shorter ones (Semjen, 1992; Van Galen, 1991; Verwey, 1995; Verwey & Dronkert, 1996). Taken together, these notions might explain the slower response times in the 333 condition as

---

[2] Part of the evidence for concurrent preparation comes from findings of slowed execution (e.g., Van Galen, 1991). Therefore, the term *concurrent* is considered more appropriate than *parallel,* which suggests interference-free processing. The data do not allow distinguishing processes involved in preparation. Thus, preparation is seen as comprising any process that occurs before execution of a response group and may include response selection, motor programming, and motor adjustment (Sanders, 1990).

compared with the 45 condition in that selecting information for the next response group occurs during execution of the preceding response group, which is slowed less as it is longer. This assertion needs verification.

This article examined the development of chunks and concurrent preparation in an experiment similar to the one carried out by Verwey and Dronkert (1996). The Structured condition was practiced extensively and involved a 333 condition with 3 three-key groups separated by three variable RSIs, and a 36 condition with a three- and a six-key group separated by two variable RSIs (the first RSI always preceded the nine-key sequence). To investigate chunk development, some blocks of trials entailed Unstructured sequences in which all RSIs had been removed. The notions expressed here gave rise to four hypotheses. With respect to chunk development it was expected that, first, response times in the Unstructured condition would reflect the positions of the RSIs in the Structured condition. This would replicate Verwey and Dronkert's results. Second, response times in Unstructured sequences should reach their asymptotic values much slower than Structured sequences because performance in the Structured condition would rely largely on the general buffer loading mechanism, whereas performance in the Unstructured condition would depend on the existence of motor chunks (which, of course, developed in the Structured condition). With extensive practice, however, Unstructured response times could approach the level of Structured response times again. Third, the differences amongst within-group response times in the relatively long six-key groups, which were expected to emerge as a result of reloading the buffer, would gradually reduce, as chunk development would make reloading superfluous. This would hold in both the Structured and the Unstructured conditions. Fourth, with respect to concurrent preparation it was anticipated that six-key response groups would be slowed less than three-key response groups—when comparing Structured and Unstructured conditions—whereas slowing of three-key response groups would be independent of whether they are part of the 333 or the 36 sequence.

In an attempt to reduce individual differences and to strengthen practice effects, this study involved some procedural changes as compared with the Verwey and Dronkert (1996) study. First, the participants had about 50% more practice than those in the Verwey and Dronkert study. Second, because intervals of a fixed duration (i.e., rhythm) might become part of a sequence representation (Keele & Summers, 1976), response groups were separated by a variable rather than a fixed RSI. It was hoped that, during practice, this would increase the tendency to prepare the response groups separately. The RSIs used were assumed to be sufficiently long for advance preparation of the response groups. Third, subjects were explicitly instructed to use the RSIs in the Structured condition for preparing the forthcoming responses. It was expected that these procedural changes would lead to higher ratios of group-start to within-group response times and to more homogeneous performance across participants than in the Verwey and Dronkert study.

Method

Tasks. A block started with the instruction on the screen to position the left little, ring, middle, and index fingers on the z, d, f, and g keys of an ordinary PC keyboard and the right thumb, index, middle, ring, and little finger on the space bar, j, k, l, and / keys, respectively. These assignments were chosen so that each finger could easily press a separate key (Figure 1). The computer screen displayed white outlines of nine squares in the same spatial arrangement as the assigned keys. The task started when the area enclosed by one of the nine squares became homogeneously green as if a light had been turned on. Participants responded by pressing the corresponding key, whereupon the green content disappeared as if the light had been turned off. After a predetermined RSI, one of the other eight squares turned green, which was again followed by pressing the corresponding key. In this way a sequence of nine keypresses was carried out in which each of the nine keys was pressed once. Keys could be released after ensuing ones had been depressed. Immediately on completion of the nine-key sequence, the next trial started, which involved the same sequence of nine keypresses.

The RSIs in the Structured condition always occurred at the same positions of the sequence and had a variable duration. To prevent participants from anticipating the moment of stimulus arrival, this variable duration interval was non-aging. Non-aging intervals are intervals with a larger probability of shorter than of longer durations (for an elaborate discussion, see Gottsdanker, Perkins, & Aftab, 1986). The interval was always in the 500-ms to 4,000-ms range. Half of the subjects performed in the 333 group. They practiced with the following time structure: NAI-0-0-NAI-0-0-NAI-0-0 ms (non-aging RSI between R_9 and S_1, 0 ms between R_1 and S_2 and between R_2 and S_3, etc.). The remaining subjects performed in the 36 condition and practiced the same keypressing sequence with the RSI sequence NAI-0-0-NAI-0-0-0-0-0 ms. The Unstructured sequences did not contain intervals between response onset and stimulus presentation.

The same basic sequence was used for all participants but each of the nine keys functioned as the starting key for two participants each of the 333 and the 36 groups. For example, when the stimulus locations are designated 1 through 8 for the fingers from left to right and the right thumb is designated 9, one sequence was |591|742|683 (i.e., |J space Z|LGD|K / F, see Figure 1). The first vertical line in this sequence indicates RSI in the 333 and 36 conditions, and the second line indicates the RSI in the 333 condition. Two other participants of each group executed 917|426|835, two performed 174|268|359, and so on. In this way, all response times had all between-hands and within-hand transitions. This is important because between- and within-hand transitions are known to affect the time between subsequent keypresses (Coover, 1923; Kornblum, 1965; Lahy, 1924).

Procedure. On Day 1, a written instruction was handed out to the participants that briefly introduced the task and the way to

```
Q    W    E    R    T    Y    U    I    O    P

  A    S    D₂   F₃   G₄   H    J₅   K₆   L₇   ;

    Z₁   X    C    V    B    N    M    ,    .    L₈

         s   p   a   c   e       b   a   r₉
```

Figure 1. Layout of the response keys on an ordinary PC keyboard. Underlined keys and the space bar were operated by nine different fingers. Indices denote location numbers that are used in the text to indicate response order.

control the computer. Participants were instructed to type as fast and as accurately as possible to maximize their score at the end of each block. They were told that the five highest scoring participants of each group of 18 would earn a bonus. All individual blocks were also preceded by a written instruction on the computer monitor, again indicating the sequence to be pressed and, in the Unstructured condition, that no RSIs would occur. At the beginning of Day 2, the experimenter explicitly instructed participants to use the RSIs for preparing the forthcoming keypresses, as the data in Verwey and Dronkert (1996) had shown large individual differences in this respect. Some participants spontaneously replied that they had not considered the possibility.

Each participant carried out seven sessions on three consecutive mornings or afternoons and one on the fourth day (the remaining Day 4 sessions are discussed in the transfer phase). Six participants performed the task simultaneously on six different computers. Three of them were in the 333 and three in the 36 group. Six other participants relaxed in an adjacent room. After a session, the first group of 6 participants could relax and the second group performed the task. This resulted in a rest and test schedule of about 17 min for each participant. Given the total of 36 participants, the experiment required the presence of 3 × 12 participants at the institute.

All sessions consisted of four blocks of trials. The fourth block of Sessions 1, 3, 5, 7, 9, 11, 13, 16, 18, 20, and 22 had the Unstructured condition.[3] The fourth block of the other sessions involved the Structured condition. Each block had 30 trials, and blocks were separated by a 20-s break.

In the practice phase, each participant pressed the nine keys in a fixed order. Hence, the participants soon knew which key to press next. In Structured sequences this had the effect that keys could be pressed before the RSI had elapsed and the stimulus had been presented. When this happened, a "too early" message was presented. An error message also occurred when an incorrect key was pressed or when no key was pressed at all during a 3500-ms interval. In these three situations, keypressing could only continue after the correct key had been pressed.

Each block was followed by display of a score that ranged from 0 to 100 points. The score consisted of a weighed combination of speed and accuracy. Given that performance improvement obeys a power law (Newell & Rosenbloom, 1981), the score was determined with a logarithmic function so that late in practice a relatively small improvement still yielded a perceivable score increase. Accuracy affected scoring in that high error rates were "punished" by reducing the score: Each additional percentage error equaled 20 ms slower responding. Error rates over 6% elicited the instruction to reduce errors. To prevent cautious and, hence, slow keypressing, error rates of less than 3% evoked the instruction to increase keying speed, unless the average response time was below 150 ms. Below 3% errors, the error rate was artificially increased before calculating the score. Thus, with 3% errors, the RT-based score was reduced least. Participants were not informed about this procedure, and average response times of error scores were not displayed.

Apparatus. The experiment was conducted on six identical IBM AT-compatible (386) computers with NEC Multisync VGA 3D color monitors. Stimulus presentation and response registration were controlled through Micro Experimental Laboratory software (MEL; Schneider, 1988). This software package is specially developed for running PC-based experiments. At a typical viewing distance of about 65 cm, a square subtended a visual angle of approximately 1°. The stimuli consisted of a bright green area filling the outline of a bright white square on a black background and were viewed under normal room illumination. The response

keys were part of a normal AT-like keyboard (BTC). Although MEL can measure times with 1-ms precision by reprogramming the internal timer, variances caused by keyboard delays were found to add approximately 19 ms to the error variance which, given the large number of trials in this study, is considered acceptable (Segalowitz & Graves, 1990).

Six participants were simultaneously tested in six sound-attenuated 2.4 × 2.5 × 2 m rooms. They sat in front of a table on which the keyboard and a computer monitor were positioned. The experimenter monitored participants through a closed-circuit video.

Participants. Participants were 36 paid students (15 men and 21 women) from the University of Utrecht. Eighteen participants were randomly assigned to the 333 and 18 to the 36 group. They were paid 180 Dutch guilders for participation. Five participants in each group received a bonus of 50 guilders. Four participants of the 36 group were replaced because, at the three final practice sessions, their error percentage exceeded 10%.[4]

Design. Analyses of variance (ANOVAs) were carried out on mean response times per condition, subject, and location of the keypress in the sequence. Response times—the time between onset of a stimulus and depression of the corresponding key—were designated T_1–T_9. In the 333 group T_1, T_4, and T_7 were termed *group-start response times* and the remaining response times were *within-group response times*. In the 36 group, T_1 and T_4 were called group-start response times and the remaining response times were within-group response times.

Keypresses involving an error and the two keypresses following that error were discarded from analysis. In each block, the first two trials were considered as warm ups and also discarded. To eliminate outliers, the 2% longest values were excluded from the analyses. To obtain independence of means and variances, we carried out arcsine transformations on mean error rates per cell before we subjected the data to ANOVAs (Winer, Brown, & Michels, 1991). Comparing three-key groups in the 333 and 36 groups required analyses, with 333 versus 36 as the between-subjects variable. Because there is no reason to expect differences between the 3 three-key response groups in the 333 condition, T_1, T_4, and T_7 were pooled in the analyses. Likewise, T_2, T_5, and T_8 as well as T_3, T_6, and T_9 were pooled. Throughout the remainder of this article, the three-key group in the 36 condition is denoted by 36-3. Besides the between-subjects ANOVA on 333 and 36-3, between-subjects ANOVAs were carried out on 333 and the six-key group in 36 (i.e., 36-6) and within-subject ANOVAs on 36-3 versus 36-6. Session covered the effect of practice, which was significant as a main effect in all ANOVAs on response times (all $ps < .001$) and is not reported separately. All ANOVAs involved Key order as a variable to account for effects caused by balancing response locations across fingers. Because there were only two participants for each Key-order condition, which probably means that Key-order effects and interactions mainly reflect individual differences, Key-order effects are not reported either.

Results

Unstructured group-start versus within-group response times. The data confirmed the earlier findings that response times in the Unstructured condition reflect the pat-

_____

[3] On Day 2, one block was discarded because of time pressure. This caused the jump from Session 13 to Session 16.

[4] One of these participants later said that he was dyslexic and had discontinued a typing course twice because of poor results.

tern of RSIs in the Structured condition (Figure 2). In the Unstructured condition, group-start response times were longer than pooled within-group response times: 333 versus 36-3 condition, $F(1, 17) = 61.8$; 333 versus 36-6 condition, $F(1, 17) = 96.4$, $ps < .001$, and this difference increased with practice: 333 versus 36-3 condition, Session 1, 118 ms, Session 22, 160 ms, $F(10, 170) = 2.1$, $p < .05$; 333 versus 36-6 condition, Session 1, 107, ms, Session 22, 170 ms, $F(10, 170) = 3.7$, $p < .001$.

Practice effects in Structured and Unstructured conditions. Overall, group start response times were longer in the Unstructured than in the Structured condition: 333 versus 36-3, $F(1, 17) = 19.2$; 36-6, $F(1, 9) = 12.4$, $ps < .001$. As shown in Figure 3, group-start response times dropped faster with practice in the Structured than in the Unstructured condition, but later the disadvantage for the Unstructured condition reduced again because the response time reduction leveled off rapidly in the Structured and not in the Unstructured condition: 333 versus 36-3, $F(10, 170) = 8.6$; 36-6, $F(10, 90) = 16.6$, $ps < .001$. Separate ANOVAs on Sessions 1 and 3 and on Sessions 3–22 confirmed this. The faster decrease of Structured group-start times in Sessions 1 and 3 was supported by a Structure × Session interaction: 333 and 36-3, $F(1, 18) = 33.1$; 36-6, $F(1, 9) = 24.6$, $ps < .001$, as was the slower decrease in Sessions 3–22, $F(9, 153) = 6.4$; $F(9, 81) = 18.3$, $ps < .001$.

Likewise, pooled within-group response times in the Structured condition were generally slower in the Unstructured than in the Structured condition: 333 versus 36-3, $F(1, 18) = 136.4$; 36-6, $F(1, 9) = 43.3$, $ps < .001$. As shown in Figure 3, within-group response times dropped sharply after Session 1 and were later approached again by those in the Unstructured condition: 333 and 36-3, $F(10, 170) = 28.2$; 36-6, $F(10, 90) = 43.9$, $ps < .001$. As with group-start response times, the faster decrease of Structured within-group response times in Sessions 1 and 3 and the slower decrease in Sessions 3–22 was supported by Structure × Session interactions: Sessions 1 and 3, 333 and 36-3, $F(1,$
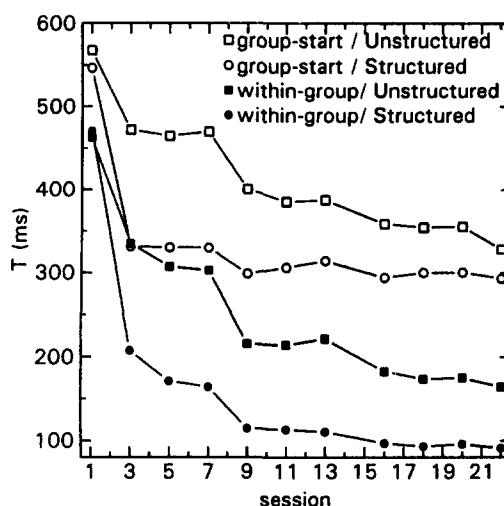

Figure 3. Group-start and within-group response times, pooled over 333 and 36 conditions, as a function of Structure and Practice.

$18) = 126.3$, 36-6, $F(1, 9) = 43.8$; Sessions 3–22, 333 and 36-3, $F(9, 162) = 12.2$, 36-6, $F(9, 81) = 22.3$, all $ps < .001$.

Individual responses within the six- and three-key groups. An ANOVA on the various Structured within-group response times in the 36-6 condition showed a Response × Session interaction, $F(80, 720) = 1.9$, $p < .001$. It was caused in part by a relatively short T_5 as compared with T_6, T_7, and T_8 in the earlier sessions: Sessions 1–4: $F(1, 9) = 32.0$, $p < .001$, which effect disappeared with practice: Sessions 19–22, $F(1, 9) = 0.9$, $p > .20$. The occurrence of some relatively long Ts in Structured within-group response times in the six-key groups of 36-6 early in practice was confirmed by the finding that pooled within-group response times in Structured 36-3 decreased more rapidly and asymptoted earlier than those in the Structured 36-6 condition, $F(21, 189) = 3.6$, $p < .001$. In addition, a main effect of
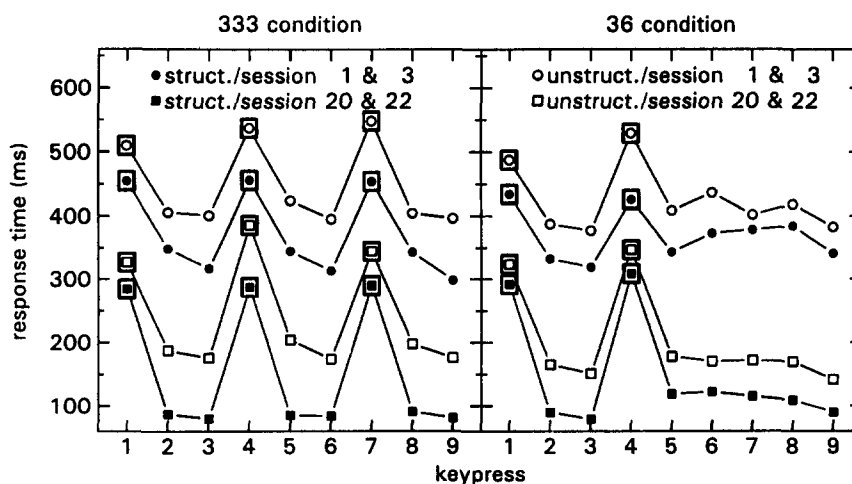

Figure 2. Response times early and late in practice as a function of 333 versus 36 condition and structure. Squares indicate the positions of response stimulus intervals in the Structured condition.

Response indicated a persistent difference between individual responses, $F(4, 36) = 3.5, p < .01$, which was caused by the relatively small T_9 as compared with T_5-T_8 (see Figure 2); planned comparison of T_5-T_8 versus T_9 on all sessions: $F(1, 9) = 15.7$; Sessions 19–22: $F(1, 9) = 9.7, ps < .01$.

In the Unstructured condition, post-hoc Tukey testing showed that the initially slow R_6 (in Sessions 1, 3, 5, and 7, $ps < .01$) became faster with practice so that all differences between R_5 and R_8 had disappeared in later sessions (in Sessions 16, 18, 20, and 22, $ps > .20$). A planned comparison to test whether the last response in 36-6 (R_9) was faster than earlier ones (R_5-R_8) confirmed this: 23 ms over all sessions, $F(1, 9) = 6.9, p < .05$. This effect was maintained during practice (Sessions 1 and 3: 33 ms; Sessions 20 and 22: 31 ms). Thus, the last response in the structured and unstructured six-key groups was and remained faster than all earlier responses, whereas the other differences amongst within-group intervals disappeared with practice.

Detailed analysis of the Structured three-key groups in 333 and 36-3 showed that the third response was generally smaller than the second one: 142 versus 127 ms, $F(1, 18) = 11.0, p < .01$. This difference disappeared with practice in 333 but not in 36-3, $F(21, 378) = 2.0, p < .01$. In the Unstructured three-key groups the difference between the second and third response did not reach significance: for 333, 23 ms, $F(1, 8) = 4.1, p = .08$; for 36-3, 14 ms, $F(1, 9) = 1.1, p > .20$.

Slowing of Unstructured within-group response times. In addition to the finding that within-group response times were slower in Unstructured than in Structured conditions, the data in Figure 2 also show that the slowing of Unstructured within-group response times was greater for three-key groups than for six-key groups. In the three-key groups (i.e., in 333 and 36-3), Unstructured response times were 104 ms longer than Structured response times, whereas in the six-key group (36-6), this was only 72 ms: 333 versus 36-6, $F(1, 18) = 8.0, p < .01$; 36-3 versus 36-6, $F(1, 9) = 7.8, p < .05$.

Ratios. The ratio of group-initiation to within-group in the Unstructured condition showed an increase with practice. They all started at about 1.0 in Session 1 and increased to 2.3 (333), 2.4 (36-3), and 2.2 (36-6) in Session 22: 333 versus 36-3, $F(10, 180) = 7.9$; 333 versus 36-6, $F(10, 180) = 10.4$; 36-3 versus 36-6, $F(10, 90) = 5.8$, all $ps < .001$. The 95% confidence interval in Session 22 ranged from 1.5 to 3.2 in the 333 condition, from 1.6 to 3.2 in the 36-3 condition, and from 1.8 to 2.7 in the 36-6 condition. Thus, the ratios found in the last session did not significantly exceed the criterion ratio of 2.0. The increase with practice did not appear to differ for the various response groups: 333 versus 36-3, $F(10, 180) = 0.9$; 333 versus 36-6, $F(10, 180) = 0.8$; 36-3 versus 36-6, $F(10, 90) = 1.4$, all $ps > .10$.

Individual differences. In the last Unstructured session (Session 22), individual ratios ranged between 0.8 and 8.0 in 333 (mean $= 2.3$), between 0.5 and 7.4 in 36-3 (mean $= 2.4$), and between 1.1 and 4.9 in 36-6 (mean $= 2.2$). The correlation between individual ratios and individual within-group response times in the Unstructured condition was highest early in practice: Sessions 1–8, 333, $r = -.72$;

36-3, $r = -.75$; 36-6, $r = -.70$, $ps < .001$, and reduced with practice: Sessions 16–22, 333, $r = -.50$, $r < .05$; 36-3, $r = -.38$, $p > .10$; 36-6, $r = -.32$, $p > .20$. In contrast, the correlation between individual ratios and individual group-start response times was not significant early in practice: Sessions 1–8, rs between .10 and .34, $ps > .16$; however, it increased with practice: Sessions 16–22, 333, $r = .66$; 36-3, $r = .66$; 36-6, $r = .57$, $ps < .01$. These correlations demonstrated that participants who had a strong tendency to group responses early in practice (i.e., large ratios) in Unstructured trials had relatively small within-group response times in those trials but not long group-start response times. Later in practice, participants with high ratios in the Unstructured condition had relatively long group-start response times and the small within-group response times were less pronounced as compared with participants with smaller ratios.

Correlations between the average of all response times in the Unstructured sequences for each participant (i.e., a general performance index) and the individual ratios of the Unstructured condition showed a relationship between general performance and grouping in Sessions 1–7: 333 and 36-3, $r = -.37$; 36-6, $r = -.56, ps < .05$, which decreased considerably in Sessions 16–22: 333 and 36-3, $r = .01$; 36-6, $r = -.13; ps > .20$. This suggests that sequences were performed better when grouping was more prominent in early practice but that this relationship reduced with practice. The correlations between average ratios in Sessions 1–7 and in Sessions 16–22: 333, $r = .57$; 36-3, $r = .65$; 36-6, $r = .60$, $ps < .01$, suggest that participants who had grouped the Unstructured sequence more in Sessions 1–7 also grouped more in Sessions 16–22.

Some additional findings: effects of response group length. An effect of response group length emerged in the Structured within-group response times. The difference between pooled within-group response times in the three- and six-key groups was marginally significant in the 333 versus 36-6 ANOVA: 333, 139 ms; 36-6, 169 ms, $F(1, 18) = 3.2$, $p = .08$, and highly significant in the 36-3 versus 36-6 ANOVA: 36-3, 130 ms, $F(1, 9) = 45.9, p < .001$. However, Structured group-start response times in six-key groups were not longer than those in three-key groups: 333 versus 36-3, $F(1, 18) = 0.8$, 333 versus 36-6, $F(1, 18) = 0.9$; 36-3 versus 36-6, $F(8, 9) = 0.8$, all $ps > 20$.

Comparison of the Unstructured group-start response times did show a significant difference between 36-3 and 36-6, $F(1, 9) = 15.9, p < .01$. This effect of sequence length on group-start intervals (i.e., the complexity effect) changed with practice, $F(10, 90) = 2.8, p < .01$: In Session 1 the difference amounted to 3 ms (T_1: 553 ms; T_4: 556 ms), in Sessions 3 and 5 it increased up to 80 and 90 ms (Session 3 T_1: 422 ms; T_4: 502 ms; Session 5 T_1: 402 ms; T_4: 492 ms), and then it gradually reduced again to 22 ms in Session 22 (T_1: 310 ms; T_4: 332 ms). Both the increase of the complexity effect in Sessions 1, 3, and 5, and its subsequent reduction were significant, $F(2, 18) = 5.1$, and $F(8, 72) = 2.6$, respectively; $ps < .05$. Comparison of the complexity effect when comparing 333 and 36-6 in the Unstructured condition did not reach significance, $F(1, 17) = 0.3, p >$

.20. When comparing pooled within-group response times in the Unstructured response groups of 36-3 and 36-6, there was a trend toward longer within-group response times in 36-6, $F(1, 9) = 3.9, p = .08$.

In short, a complexity effect was found with respect to Unstructured three- and six-key response groups in the 36 condition, and a trend was found that 36-6 response groups were executed more slowly than 36-3 response groups. Here, the size of the complexity effect first increased and later decreased again. No complexity effect was encountered in the Structured condition but, here, six-key response groups were executed more slowly than three-key response groups.

Errors. Table 1 shows average error percentages per day in the Structured three- and six-key response groups. In the Structured three-key groups of the 333 and 36-3 conditions, error rate of the second response was found to increase on Days 1 and 2, which was not the case with respect to the first and third response, $F(42, 756) = 5.9, p < .001$. A similar effect of practice was found in the six-key response group of 36-6, $F(20, 180) = 7.2, p < .001$. Here the error increase with practice was more pronounced for later responses, with the exception of the last response, $F(100, 900) = 2.2, p < .001$. Finally, comparison of errors in Structured and Unstructured three-key groups in 333 and 36-3 showed more errors in Unstructured than in Structured conditions: Structured, 4.3%; Unstructured, 5.3%; $F(1, 8) = 112.3, p < .001$.

Discussion

The practice phase was concerned with the effects of extensive practice in a nine-key pressing sequence through which subjects cycled continuously so that one sequence was immediately followed by the next one. The first issue was whether motor chunks would develop in the Structured condition. This notion was corroborated by various aspects of the data. First, Unstructured response times clearly and increasingly reflected the positions of the RSIs in Structured conditions. This also showed as increasing ratios between group-start and within-group response times and replicated Verwey and Dronkert's (1996) findings. In addition, the present data replicated Verwey and Dronkert's findings that

the ratios obtained at the final sessions of the practice phase exceeded 2:1. These findings corroborate Verwey and Dronkert's claim that the pattern of Unstructured response times had not been caused by the learning of interkey timing but, rather, by the development of motor chunks. Notice that the fact that Structured group-start response times were longer than Structured within-group response times can be attributed solely to the present use of RSIs with random duration. Verwey and Dronkert used fixed duration RSIs and did not find longer group-start response times in the Structured condition. The second indication that motor chunks had developed in the Structured condition was that Unstructured group-start and within-group response times decreased more slowly with practice than their Structured counterparts but that, eventually, Unstructured response times approached Structured response times again. Third, early in practice there were differences amongst within-group response times for the six-key groups, which disappeared later in practice (with the exception of the last response time which remained fastest; see later for an explanation). This effect was observed in the Structured as well as the Unstructured conditions and is in line with the notion that motor chunks dominate the execution of longer response groups in the Structured condition as well. (Later, an alternative explanation is examined.)

The second issue in the practice phase concerned preparation of a forthcoming response group during execution of its predecessor. This notion of concurrent preparation is supported by the finding of longer within-group response times in Unstructured than in Structured response groups, which had been observed before by Verwey and Dronkert (1996). Thus, preparation processes and execution processes appear to concur and share a limited processing capacity. As anticipated, lengthening of the Unstructured within-group response times was greater for three- than for six-key groups, whereas it did not depend on the length of the forthcoming response group. This corroborates the hypothesis advanced in the introduction stating that the amount of processing required for preparing a response group for which a motor chunk exists is independent of the size of that response group, whereas concurrent preparation slows execution of the ongoing response group less as this group is longer.

The ranges of group-start ratios between and within-group response times observed with individual participants in the various conditions did not appear very different from those reported by Verwey and Dronkert (1996) and, in fact, the mean ratios in the final sessions of the present study were even slightly smaller. This indicates that the procedural changes in the present study—more practice, variable RSIs, and instructions to prepare during RSIs in the Structured condition—did not affect the way the Unstructured sequences were carried out.

One important aspect of these data, which was also observed by Verwey and Dronkert (1996), is that in the Unstructured block of the last session some subjects still had group-start to within-group ratios near 1:1; in Session 22, 7 out of the 18 (333 condition) participants and 7 out of the 18 (36 condition) participants had ratios below 1.5.

Table 1
Error Percentages in the Structured Three- and Six-Key Response Groups As a Function of Day

Day	Response in three-key group			Response in six-key group					
	1	2	3	1	2	3	4	5	6
1	4.3	3.9	4.1	3.6	3.2	5.3	5.1	5.3	4.4
2	4.9	6.2	3.2	3.7	3.8	6.3	7.4	10.1	4.7
3	5.0	6.4	3.0	3.3	4.6	6.0	7.1	9.8	3.8
M	4.7	5.5	3.4	3.5	3.9	5.9	6.5	8.4	4.3

Note. Day 3 includes Session 22 from Day 4 as well.

Thus, whereas the greater proportion of the participants clearly partitioned the Unstructured sequences, some participants had apparently produced the Unstructured sequence with little reliance on sequence partitioning. These large individual differences were probably the reason that the average group-start to within-group ratios in the last session did not significantly exceed the 2:1 criterion ratio. As suggested by the correlations between average response times and group-start to within-group ratios, the participants with small ratios performed poorly in the earlier sessions but could eventually reach performance levels that were comparable with those of the other participants. This observation is important, as it suggests that there are other ways of reaching high performance levels than learning the imposed chunks. Perhaps these participants had higher degrees of concurrent preparation (which reduces the ratio) or they had partitioned the Unstructured sequences in some other way (e.g., because, given their assignment of fingers to sequence elements, some participants could recall another way of partitioning more easily). The latter explanation is supported by the finding that grouping was relatively stable over practice for individual participants. Perhaps, some participants even developed a single chunk for the entire sequence. These findings illustrate that, on the one hand, future research should involve a more profound way of imposing a grouping structure. On the other hand, they suggest that different strategies in early practice may result in different chunk structures after extensive practice.

Examination of the individual within-group responses showed that the last response in Structured and Unstructured six-key groups was, and remained, faster than all earlier responses. In Structured three-key groups, this effect was also observed early in practice but it reduced in Structured 333 response groups and it did not reach significance in Unstructured three-key groups. A fast last response in a response group is not unexpected, as this has been found in earlier studies as well (Verwey, 1994; 1995). There it occurred when producing relatively short keypressing sequences with one finger; it was taken to indicate that retrieving a single response from the motor buffer concurs with execution of the preceding response. The present finding of a fast last keypress is a multifinger keypressing sequence indicates that, in the earlier studies, the effect had not been caused by some (biomechanical?) factor associated with one-finger keypressing. That concurrent preparation of an entire response group and concurrent retrieval of individual sequence elements are two independent types of concurrent processing is indicated by the fact that the last response was fastest in Structured as well as Unstructured six-key groups.

One basic assumption in this study was that motor chunks develop when, in the Structured condition, the motor buffer is consistently loaded with the individual elements included in a single response group in the way described by Henry and Rogers (1960) and Sternberg et al. (1978). That normal buffer loading took place in the present study and that the response group was programmed as a sequence is corroborated by various findings: (a) A complexity effect was found in the Unstructured 36 condition in that the group-start

interval was longer for 36-6 than for 36-3, (b) a sequence length effect was observed in Structured and Unstructured within-group response times (Sternberg et al., 1978), (c) the last response in most response groups was relatively fast (see earlier), and (d) there was a fixed distribution of errors across responses in Structured response groups. For now, it is unclear to what extent these indications for buffer usage disappear with extensive practice when a motor chunk develops. The fact that the complexity effect in the Unstructured 36 condition reduced with practice and was absent in Structured response groups, and the fact that the last response in the Structured three-key groups of the 333 condition was no longer fastest with practice and did not reach significance in Unstructured three-key groups, might indicate that, with extensive practice, the development of chunks makes these typical indications of motor buffer usage disappear. This would be in line with findings that when sequences are made up of highly practiced elements, each of these elements may themselves be a sequence of relatively simple movements (e.g., Sternberg et al., 1978: pronouncing word sequences; Thomassen & Van Galen, 1992: writing letter sequences).

Before completing this discussion, an alternative explanation for the gradual disappearance of the differences between within-group intervals in the six-key groups deserves consideration. The interprepation that a single chunk had developed for representing the six-key group contrasts with studies suggesting that longer sequences are, more or less spontaneously, divided into chunks of two elements and that longer sequences remain a concatenation of more than one chunk (i.e., forming a binary tree representation; Frensch, 1994; Gordon & Meyer, 1987; Rosenbaum et al., 1983). The present data do not directly exclude the possibility that, because of the extensive amount of practice in the present study, response times were subject to floor effects and, therefore, could not show that several chunks had been executed in rapid succession.[5] However, two aspects of the data argue against this possibility. First, the response times in the Unstructured sequences of Sessions 20 and 22 (Figure 2), which were slower than those in Structured sequences and probably not limited by floor effects, did not involve one or more relatively long response times as would be expected when a new subchunk is being loaded. Second, the error distribution across the six-key group showed a pattern comparable with the three-key group (Table 1) in that error rate increased with position within the group—with the exception of the last. Although this phenomenon is not directly predicted from a chunking point of view, it seems more in line with the notion that the six-key group was performed as a whole than that it would consist of several independent chunks. Perhaps, indications that longer sequences are divided into separate subchunks, such as those found in the earlier sessions of the practice phase (Figure 2), merely reflect a temporary state and, with additional practice, these subchunks integrate into a single one. This explanation is in line with the observation that Un-

structured group-start response times decreased more slowly in 36-6 than in 36-3.

In summary, the results in the practice phase support the position that motor chunks develop when the same response elements are repeatedly loaded into a short-term motor buffer. As such, the present data are in line with those presented by Verwey and Dronkert (1996) and refute an explanation in terms of learning a rhythm for movement sequences performed at the maximal performance rate. (Of course, the role of rhythm learning at submaximal rates cannot be disproved from the present data.) The development of motor chunks follows from the findings that (a) Unstructured response times clearly and increasingly reflected the positions of the RSIs in Structured conditions; (b) Unstructured group-start and within-group response times decreased more slowly with practice than their Structured counterparts but, eventually, Unstructured response times approached Structured response times again; and (c) differences amongst within-group response times for the Structured and Unstructured six-key groups disappeared later in practice. Besides, preparation for a forthcoming group of responses can occur during the production of an earlier response group. The amount of slowing of the ongoing response group seems to depend on the size of the ongoing response group and not on the size of the prepared response group. This corroborates the notion that the time of loading a motor chunk is virtually independent of the size of the movement sequence it represents, as is also indicated by disappearing complexity effects with practice. Finally, the use of RSIs during practice in the Structured trials may not always have had the effect that participants partitioned the sequence into the imposed response groups or used a comparable amount of concurrent preparation. Some participants may have partitioned the sequence in another way or were able to prepare a response group to a more progressed state during execution of the preceding response group.

Transfer Phase

The transfer phase addressed performance of keypressing sequences that were more or less different from the ones exercised in the practice phase. As such, it tested the contention that motor chunks are specific with respect to their content (Sternberg et al., 1990). If motor chunks are content specific, changing even a single element in a response group should have the effect that performance of the entire response group deteriorates considerably. This would accord with various models of skill outside the domain of motor learning (Allport, 1980; Estes, 1986; Logan, 1988, 1992; Welford, 1968). Evidence of content-specific effects of practice in the production of movement sequences has been mainly derived from error analyses (Drummond, 1981; Fentress, 1983; Fromkin, 1981; Gallistel, 1980; Zimmer & Körndle, 1988) but response time should be affected as well.

The notion of content-specific practice effects is at odds with experimental results showing that performance in practice and in new movement sequences do not differ (Cham-

berlin & Magill, 1992b; Verwey, 1992). Neither did practicing one particular pair of movement sequences or four variations of such a pair show a performance difference (Verwey, 1994). These findings suggest that practice has foremost content-aspecific effects, which is consistent with notions that memory representations of motor skills have a generally abstract form (a general motor program or schema) that can be easily adapted to different situations (Chamberlin & Magill, 1992a; Schmidt, 1975, 1982).

The transfer phase tested an explanation for the apparent contrast between the theoretically expected content-specific and experimentally obtained content-aspecific effects of practice. This explanation emerged from the notions pointed out before: When there is ample time for preparing the execution of a sequence, performance would rely on advance buffer loading and execution rate is relatively insensitive to previous practice. Indeed, Chamberlin and Magill (1992b) and Verwey (1992) used discrete sequence production tasks in which sequences of two to four elements were produced following ample preparation time, whereas evidence for content-specific learning effect is typically found in more realistic tasks that do not allow movement sequences to be prepared at ease (e.g., in writing: Thomassen & Van Galen, 1992; controlling a pedalo: Zimmer & Körndle, 1988). Content-specific effects of practice would emerge only when the sequence is relatively long so that new sequences cannot be programmed entirely in advance and when the time required for loading the motor buffer has an effect on overall performance, such as in the Unstructured condition of the practice phase.

This explanation was examined in two conditions that differed with respect to the opportunity participants had for advance preparation. This first condition was the Unstructured condition that had also been used in the practice phase. Because of the absence of RSIs in this condition, any advance preparation of groups of response would show up in task performance. Performance in this condition was assumed to depend heavily on the existence of motor chunks. In the Discrete condition, execution of each response group was preceded by presentation of the successive stimuli in the appropriate order. This procedure gave ample possibilities for advance preparation of a response group. For 333-condition subjects, each Discrete response group consisted of three keypresses; for 36-condition subjects, Discrete response groups included either three of six keypresses.

Both the Unstructured and the Discrete conditions contained four sequence types, each of which was produced by all participants. The first sequence type involved the Practiced response groups from the practice phase. The second sequence type included New response groups in which the order of keys was entirely unpredictable from what had been learned in the practice phase. Whereas Practiced three- and six-key response groups were expected to be rapidly produced in both the Discrete and the Unstructured conditions, New response groups would be produced rapidly only in the Discrete condition and only in case they would not exceed the motor buffer capacity. This would probably hold

for three-key groups. However, New six-key response groups were likely to exceed the motor buffer capacity and would, therefore, show some relatively long within-group response times caused by reloading the buffer during execution of the response group. Furthermore, the New sequences in the Unstructured condition were expected to be produced on a one-by-one basis: each keypress in response to the corresponding stimulus. Confirmation of these predictions would offer support for the use of motor chunks in familiar movement sequences and the heavy reliance on loading a limited capacity motor buffer in advance when new movement sequences are produced.

Two further sequence types were included for testing the specificity of motor chunks. The ABB and AAB sequence types consisted of response groups in which the response order was as practiced in the practice phase but there these response had belonged to successive response groups. These sequence types were introduced to test whether, in the practice phase, associations had developed between responses of successive response groups, just as they would have for responses of a single response group (Brown & Carr, 1989; MacKay, 1982). In the Discrete ABB and AAB conditions, 333-condition subjects performed three response groups of three responses each. In ABB, these response groups consisted of the third response of a response group in the practice phase (denoted by the A) as first response, and the first and second response of the next response group in the practice phase (denoted by BB) as second and third response. In a similar manner, in AAB the second and third response of one response group of the practice phase (denoted by AA) were followed by the first of the next response group (denoted by B). (The reader is referred to Table 2 for a more formal description of the ABB and AAB sequences.) From the chunking point of view, which assumes associations between elements within a single response group, the second response in ABB and the third in AAB should be slower than the other responses.

Subjects in the Discrete 36 condition performed, again, a three- and a six-key response group. The procedure described for rearranging responses in the 333 condition was also used for obtaining the ABB and AAB sequences in the 36 condition. This had the result that the ABB and AAB six-key groups involved parts of the Practiced groups embedded in a new context. Here, the issue was the extent that these familiar parts would improve performance in the ABB and AAB conditions relative to the New condition. Thus, the general question with respect to the ABB and AAB sequences in the 36 condition concerned the extent that motor chunks can be used in dissimilar situations. A content-specific view predicts performance in ABB and AAB to be similar to that in the New condition, whereas a content-aspecific view expects performance to be high in all conditions, irrespective of differences relative to the practiced sequences. Detailed examination of individual response times should cast light on the origins of performance differences amongst the various conditions.

These notions permit the following predictions:

1. In the Discrete condition, the differences between within-group response times in the Practiced and in the New sequences should be smaller for three-key groups than for six-key groups because Practiced and New three-key groups and Practiced six-key groups can be entirely loaded into the motor buffer in advance, whereas New six-key groups need reloading at some point.

2. In the Unstructured condition, response times in the Practiced condition should reflect the pattern of RSIs of the practice phase, whereas in the New condition individual keys are pressed in response to the corresponding stimuli and no grouping is expected.

3. On the assumption that the occurrence of familiar sequence parts in otherwise new contexts does not allow the use of motor chunks, it can be assumed that performance in ABB and AAB response groups is comparable with performance in New response groups. This should hold for Discrete and for Unstructured conditions.

Method

Tasks. Participants who had served in the practice phase also were involved in the transfer phase. They carried out keypressing sequences in two Timing conditions: the Discrete and the Unstructured conditions. In the Discrete condition, each trial consisted of two (in the 36 condition) or three (in the 333 condition) groups of sequential keypresses. Before each response group was produced, the computer informed participants about the responses by sequentially filling either three or six squares with a white content and presenting a number at the center of each square indicating its location within the response group. The rationale behind this procedure was that new response groups could be produced with ample opportunity for advance preparation. Onset asynchronies between squares in this instruction part of the Discrete condition amounted to 400 ms. After the last square had been lit, the display remained unchanged for 2,500 ms, showing all filled squares with their location number. Then all squares were cleared and after a non-aging interval between 500 and 4,000 ms, the first square in the response group turned green, indicating the stimulus corresponding to the first keypress in the response group. Immediately after the corresponding key was pressed, the square was cleared and the next one in the response group turned green. This was repeated until all three or six keys had been pressed. Thus, there was only temporal uncertainty at the start of the response group but no uncertainty with regard to the keys in the group and their order. When an error was made or when an additional key was pressed within 500 ms after pressing the last key of the response group, an error message was presented. Next, the ensuing response group was shown and participants were asked to repeat it. The order of the keys within the two of three response groups in a trial of nine keypresses remained the same in one block of trials.

The response groups in the Discrete condition were identical and in the same order as in the *Unstructured* condition. However, in the Unstructured condition of the transfer phase, each stimulus followed the preceding response immediately (no RSI).

Each of these two Timing conditions involved the same set of four sequence types: one as practiced in the practice phase, one new, and two in which parts of the practiced response groups occurred. These sequences appear in Table 2. This table shows that in the New condition, each response followed another response than in the practice phase with the exception of the last two

Table 2

Overview of the Sequences Performed in the Transfer Phase for Participants in the 333 and 36 Conditions

Sequence	Condition	
	333	36
Discrete		
Practiced	123|456|789	123|456789
New	765|198|423	765|198423
AAB	234|891|567	234|891567
ABB	345|912|678	345|912678
Unstructured		
Practiced	123456789	123456789
New	765198423	765198423
AAB	234891567	234891567
ABB	345912678	345912678

Note. The numbers indicate the positions of the keys as pressed in the practice phase and in the Practiced sequences of the Transfer phase. As described in the practice phase, the actual order differed for individual participants. Vertical lines in the Discrete condition represent the start of separate response groups.

responses.[6] The two sequence types containing parts from the practiced sequences—AAB and ABB—are characterized by the fact that all three-key response groups in the 333 condition involved keys that had also followed each other during practice. Either the first and second (in AAB) or the second and third (in ABB) had originally occurred together in a single response group in the practice phase, whereas the remaining third (in AAB) or first (in ABB) key had also preceded or followed the other two but had belonged to another response group. As in the practice phase, the sequences cycled across the 18 participants in a group so that there were pairs of participants in the 333 and the 36 conditions having the same sequences.

A session in the transfer phase contained either the Discrete or the Unstructured condition. Half of the participants performed a Discrete session as first and third sessions and an Unstructured session as second and fourth sessions. For the remaining participants this was reversed. Each participant performed the four sequence types in a single session. The order of sequence type blocks in each session was balanced over participants.

Procedure and design. After 3 days of practice in the practice phase, participants performed five sessions at Day 4. The first session has been reported as Session 22 of the practice phase. Then, before the four transfer sessions, participants received a brief introduction describing the transfer tasks and mentioning the fact that one condition involved the practiced sequence and three were different from those that had been practiced.

Half of the 333-condition and half of the 36-condition participants started with a 17-min session containing the unstructured transfer condition. After completing this session, the participants rested and the other half of the participants started with a 30-min Discrete session. Then, the first group did the 30-min Discrete session, followed by the second group who did their 17-min Unstructured transfer session. Then, each participant repeated these two sessions in the same order.

The same equipment was used as in the practice phase. Data analyses were similar to those described in the practice phase: 333 versus 36 condition was a between-subjects variable, and sequence type, Timing, Session, and Key-order were within-subjects variables. In addition, individual response times within a response group were compared by including Response in the design.

Results

An overview of the data is presented in Figures 4 and 5. In short, Discrete three-key groups in all sequence type conditions were performed very quickly, with limited differences between the Practiced, New, ABB, and AAB sequences. In contrast, six-key groups in the New, ABB, and AAB sequences showed some relatively long within-group response times which, in the ABB and AAB sequence, tended to coincide with the first of familiar response pairs and triplets. In the Unstructured condition, performance was generally much better in the Practiced condition, where the pattern of RSIs from the Structured condition of the practice phase reappeared, than in the New, ABB, and AAB sequences. Once again, the occurrence of practiced pairs and triplets appeared to have had only a small performance effect. Thus, performance in the ABB and AAB sequences was generally more like performance in the New than in the Practiced sequence.

Discrete condition. Figure 6 gives a different representation of the data in Figures 4 and 5, in particular because within-group response times were pooled. Pooled within-group response times in three- and six-key groups were shorter in the Practiced than in the New sequences: 333 versus 36-3, $F(1, 18) = 46.8$, $p < .001$; 36-6, $F(1, 9) = 68.7$, $p < .001$. Moreover, pooled within-group response times were generally longer in six- than in three-key groups: 333 versus 36-6, $F(1, 18) = 12.6$, $p < .01$; 36-3 versus 36-6, $F(1, 9) = 76.0$, $p < .001$. According to expectations, this difference between three- and six-key response groups was larger in New than in Practiced sequences: 333 versus 36-6, $F(1, 18) = 14.6$; 36-3 versus 36-6, $F(1, 9) = 37.0$, both $ps < .001$, but remained significant when tested in the Practiced sequences alone: 333 versus 36-6, $F(1, 18) = 7.7$, $p < .001$; 36-3 versus 36-6, $F(1, 9) = 14.1$, $p < .01$.

Pooled within-group times in the ABB and AAB sequences differed from those in the Practiced sequences: 333 versus 36-3, $F(1, 18) = 77.5$; 36-6, $F(1, 9) = 119.6$, $ps < .001$. In contrast, within-group times in ABB and AAB sequences did not differ significantly from those in New sequences: 333 versus 36-3, $F(1, 18) = 0.9$; 36-6, $F(1, 9) = 2.1$, $ps > .18$.

Detailed analyses of separate within-group response times showed that in the Practiced six-key groups of 36-6, T_9 was smaller than the pooled T_5–T_8, $F(1, 9) = 19.6$, $p < .001$. This difference amounted to 35 ms and indicates concurrent retrieval of individual keypresses from the motor buffer. In a similar manner, T_9 in the 36-6 New sequences was smaller than T_5–T_8, $F(1, 9) = 26.9$, $p < .001$, suggesting grouping of R_8 and R_9. Furthermore, in the New 333-condition response groups, within-group response times of the practiced response pair (R_8 and R_9) were faster than those in the new pairs (i.e., R_2, R_3, R_5, and R_6), 18 ms, $F(1, 9) = 14.9$, $p < .01$.

In AAB, the second response of the three-key groups of

_____

[6] This was actually unintended but was nonetheless interesting to analyze.
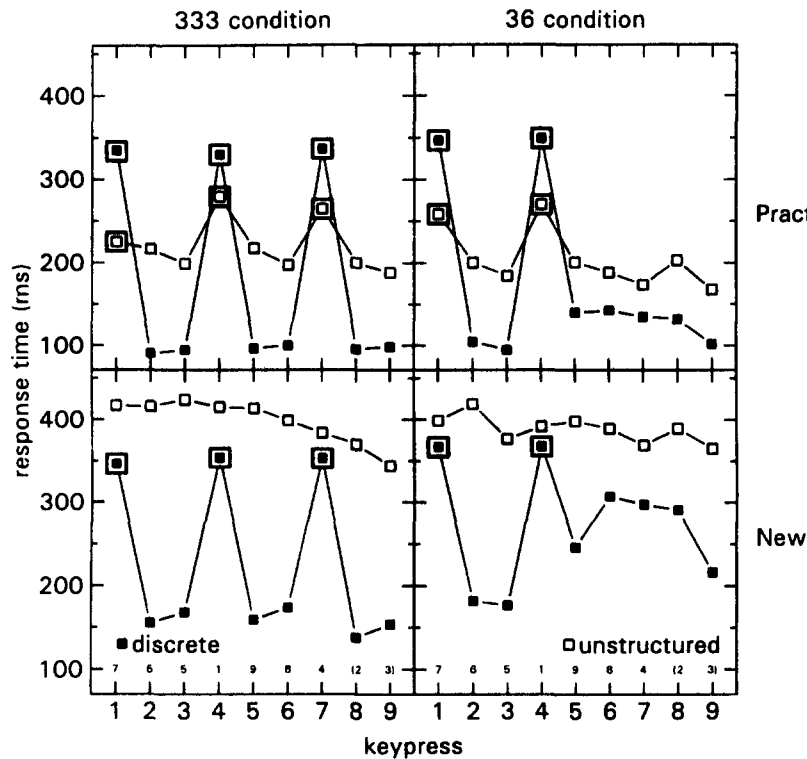
Figure 4. Response time as a function of 333 versus 36 condition, Timing, and Response in the Practiced and New sequence types. The numbers at the bottom of the New frames indicate the number of that response in the Practiced sequences. Parentheses indicate previously practiced pairs and the large squares show the start of response groups in the Discrete condition and, with regard to the Practiced condition, of the response stimulus intervals, in the practice phase.

333 and 36-3 conditions was faster than the third: 149 ms and 176 ms, $F(1, 18) = 9.7$, $p < .01$. In the ABB 36-6 response groups, T_7 was greater than T_5, T_6, T_8, and T_9, $F(1, 9) = 18.6$; $p < .001$, suggesting that R_7–R_9 formed a single response group. In AAB 36-6, both T_6 and T_7 were longer than T_5, T_8, and T_9, $F(1, 9) = 29.4$; $F(1, 9) = 22.6$, $ps < .001$, also suggesting grouping of R_7–R_9.

Error percentages associated with within-group response in the three- and six-key response groups were generally below 5%. Relatively high error scores were found in New within-group responses; averaged over three- and six-key groups: New sequences, 5.7%; Practiced, ABB, and AAB sequences, 3.7%; 333 and 36-3, $F(3, 54) = 7.4$, $p < .001$; 36-6, $F(3, 27) = 5.0$, $p < .01$.

Unstructured condition. As depicted in Figure 4, the pattern of response times in the Practiced sequences of the Unstructured condition reflects the pattern of RSIs of the Structured condition in the practice phase. Thus, within-group response times were shorter than group-start times: 333 versus 36-3, $F(1, 18) = 17.8$, $p < .01$; 36-6, $F(1, 9) = 29.6$, $p < .001$. Such a response pattern was not encountered in the New sequences: 333 versus 36-3, $F(1, 18) = 0.3$; 36-6, $F(1, 9) = 0.2$, $ps > .20$.

The response times in the ABB and AAB sequences reflected the RSI pattern from the practice phase and also

showed differences between pooled group-start and within-group times: ABB-333 versus 36-3, $F(1, 18) = 9.7$ $ps < .01$; ABB-36-6, $F(1, 9) = 9.9$, $p < .05$; AAB-333 versus 36-3, $F(1, 18) = 6.8$, $p < .01$; AAB-36-6, $F(1, 9) = 8.5$, $p < .05$. However, as depicted in the left frame of Figure 6, Unstructured group-start times were considerably shorter in the Practiced sequences than in the New, ABB, and AAB sequences: 333 versus 36-3, $F(3, 54) = 61.4$; 36-6, $F(3, 27) = 19.9$, $ps < .001$; all pairwise planned comparisons with Practiced group-start times, $ps < .001$. Likewise, pooled within-group response times were smaller in the Practiced sequences than in the New, ABB, and AAB sequences: 333 versus 36-3, $F(3, 54) = 102.9$; 36-6, $F(3, 27) = 103.4$, $ps < .001$; all planned comparisons with Practiced within-group response times: $ps < .001$. Moreover, 333 and 36-3, ABB and AAB sequences had smaller pooled within-group response times than did New sequences: New versus ABB, $F(1, 18) = 5.0$; New versus AAB; $F(1, 18) = 6.4$, both $ps < .05$. In 36-6, pooled within-group times in ABB and AAB sequences did not differ from those in New sequences: New versus ABB, $F(1, 9) = 2.3$; New versus AAB, $F(1, 9) = 3.1$, $ps > .11$. In addition, pooled within-group response times in the three- and six-key groups did not differ: 333 versus 36-3, $F(1, 18) = 0.2$; 36-3 versus 36-6, $F(1, 9) = 1.8$, $ps > .20$. Thus,
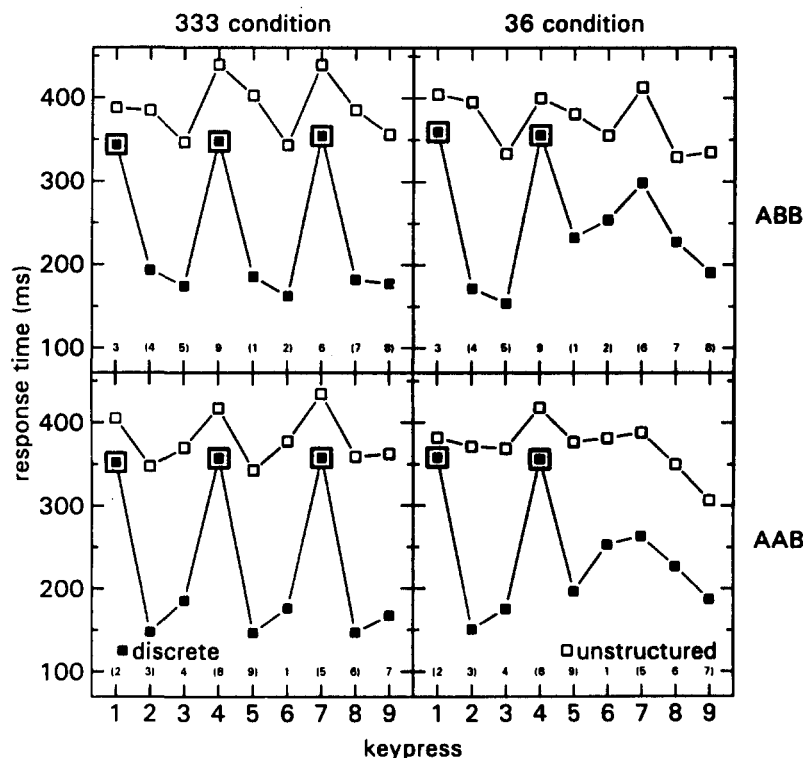
Figure 5. Response time as a function of 333 versus 36 condition, Timing, and Response in ABB and AAB. See Figure 4 for the meaning of numbers between parentheses and large squares.

despite the emergence of the RSI pattern from the practice phase in the ABB and AAB sequences, the general performance level in ABB and AAB sequences was more similar to that in New than in Practiced sequences.

Detailed examination of separate within-group responses in Unstructured three-key response groups showed a Sequence Type × Response interaction, $F(3, 54) = 8.0, p < .001$. It was caused by a faster third response in Practiced, 19 ms, $F(1, 18) = 13.5, p < .001$, and ABB, 52 ms, $F(1, 18) = 76.6, p < .001$, which effect appeared greatest in ABB, $F(1, 18) = 22.2, p < .001$. In New three-key groups, T_8 and T_9 were shorter than $T_2, T_3, T_5,$ and $T_6, F(1, 9) = 10.5, p < .01$. ($R_1, R_4,$ and R_7 were considered group-start response times and not included in this analysis.) In the ABB and AAB six-key groups, R_7 was slower than R_8 and R_9, $F(1, 9) = 21.8$; $F(1, 9) = 17.6$, $ps < .01$. Hence, familiar response pairs or triplets in New, ABB, and AAB sequences (see parentheses in Figure 5) were generally characterized by relatively fast second and third responses. Planned comparison of R_5–R_8 with R_9 in the Practiced six-key group showed a marginally significant difference: 24 ms, $F(1, 9) = 4.3, p = 0.07$.

Sequence type main effects in 333 versus 36-3 and in 36-6 on transformed error proportions, $F(3, 54) = 6.8$; $F(1, 9) = 14.9, ps < .001$, were caused by less errors in Practiced than in New, ABB, and AAB sequences (all planned comparisons with Practiced, $ps < .05$). Error percentages amounted to 3.6, 6.5, 7.7, and 5.8 in three-key groups and 2.3, 7.4, 5.6, and 6.8 in the six-key groups.

Some additional findings. the difference between Unstructured and Discrete within-group times in the Practiced sequences was greater for three- than for six-key groups, confirming similar findings from the practice phase: 333 versus 36-6, $F(1, 18) = 10.0, p < .01$; 36-3 versus 36-6, $F(1, 9) = 24.0, p < .001$ (see Figure 4).

More errors had been made in the Unstructured than in the Discrete conditions: 333 versus 36-3, 5.0% versus 3.3%, $F(1, 18) = 42.8, p < .001$; 36-6, 6.1% versus 3.6%, $F(1, 2) = 23.5, p < .05$. These effects were primarily caused by the relatively low error proportions associated with Discrete
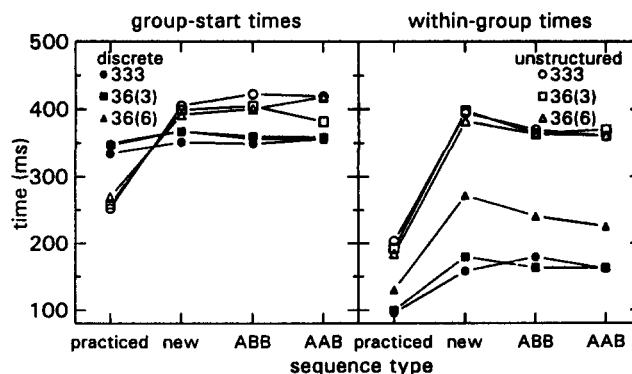

Figure 6. Group-start and within-group times as a function of Timing, Response group, and sequence type.

group-start responses (2.2%) as compared with Unstructured group-start errors (4.8%) and within-group errors (Discrete: 4.8%, Unstructured: 4.9%): 333 versus 36-3, $F(3, 54) = 4.1, p < .01$; 36-6, $F(1, 2) = 25.3, p < .05$.

Individual differences. To test whether participants who had probably relied more on grouping during practice also showed more evidence for chunk usage in the transfer conditions, correlations were computed between group-start and within-group ratios for individual participants in the Unstructured sequences of the practice phase and the ratios of response times in the transfer conditions that were supposed to be affected as participants relied more on existing chunks. More specifically, the occurrence of familiar pairs of triplets in the otherwise new sequences was expected to have the effect that the second (and third) response would be relatively fast, causing positive correlations between group-start and within-group ratios in the practice phase and the ratios between the first and later response times of a familiar group in the transfer conditions.

Table 3 shows the correlations between response times of familiar pairs and triplets in New, ABB, and AAB. In the Unstructured 333 condition, the correlations are in the predicted direction: The positive correlations show that the second response of familiar pairs was relatively fast as compared with the first response for participants who had also grouped more during the Unstructured trials of the practice phase. In the Discrete 333 condition, this seemed not to be the case, which is in line with the notion that chunks do not affect performance with short sequences when there is ample time for advance preparation.

In the 36 condition, things looked differently. With one exception, there was little evidence for the usage of chunks of familiar pairs or triplets, irrespective of whether the original chunks had been part of the three-key group during the practice phase ($R_5 R_6$ in ABB, $R_1 R_2$ in AAB) or of the six-key group ($R_2 R_3$ and $R_7 R_8 R_9$ in ABB, $R_4 R_5$ and $R_7 R_8 R_9$ in AAB). The one exception concerned the correlation between the group-start to within-group ratio in the practice phase and the R_8 to R_9 ratio in the Unstructured and Discrete New conditions. However, in contrast to expectations this correlation was negative. T_8 was much longer than T_9 in all cases (Figure 4). However, it appears that 36 condition participants who had grouped the three-key group

in the 36 confition sequence more during practice had a longer T_9 than participants who had grouped less. A post hoc explanation is that participants who had less developed chunks still knew that R_9 would follow R_8 and tended to group R_8 and R_9 explicitly by preparing them as a group (slow R_8, fast R_9), whereas in the presence of stronger R_8–R_9 associations, less explicit preparation was used, resulting in a relatively fast R_8.

Discussion

The main goal of the transfer phase was to investigate performance effects of practiced and three new response groups of differing lengths, with and without time for advance preparation. As evidenced in Figure 6, the results are entirely in line with the hypotheses expressed in the introduction to the transfer phase:

1. In the Discrete condition, the difference between pooled Practiced and New within-group response times was smaller for three- than for six-key groups. This difference was largely due to a few relatively long within-group response times in the New six-key groups (Figure 4). This confirms that the existence of motor chunks is insignificant for short response groups when there is ample time for advance preparation. However, when the response group is longer or the buffer capacity is exceeded, unfamiliar response groups need to be executed in parts, causing some relatively long within-group intervals, whereas familiar response groups can still be produced as a whole because they are entirely loaded in advance. The finding that practice has a minor effect in Discrete conditions with short sequences suggests that content-aspecific effects of practice in the production of short movement sequences can be attributed to advance buffer loading (e.g., Chamberlin & Magill, 1992a; Verwey, 1992, 1994).

2. Performance in the Unstructured condition was generally much better with Practiced than with New sequences. The return of the RSI pattern from the practice phase in the response times of the Practiced sequences of the transfer phase indicates that the motor chunks that had developed in the practice phase were used in the Unstructured condition

Table 3

Correlations Between the Amount of Grouping in the Unstructured Conditions of the Practice Phase (Indicated by the Group-Start to Within-Group Ratio) Across Individual Participants and Ratios

	333 condition			36 condition		
Sequence	Ratio	Unstructured	Discrete	Ratio	Unstructured	Discrete
New	R8/R9	.34	−.21	R8/R9	−.48*	−.50*
ABB	R258/R369	.45†	−.03	R2/R3	.06	.29
				R5/R6	.02	−.30
				R7/R8R9	.05	−.02
AAB	R147/R258	.50*	.21	R1/R2	−.09	−.17
				R4/R5	−.11	.20
				R7/R8R9	−.25	.06

† $p < .10$. * $p < .05$.

for loading the motor buffer. Apart from R_8 and R_9, the New sequences displayed no signs of grouping, which suggest that sequence production relied on the sequential selection and execution of individual responses. This difference between Practiced and New sequences supports the notion that the availability of motor chunks becomes important when there is no time for advance loading of the motor buffer.

3. The overall level of performance in the ABB and AAB sequences was more similar to performance in the New than in the Practiced sequences (Figure 6). This holds for Discrete and Structured conditions and confirms that the occurrence of familiar parts in the otherwise new response groups does not increase sequence execution rate much. Further support for this position comes from a detailed examination of the individual responses in the three-key groups of ABB and AAB, which showed that the second of a familiar response pair was only slightly faster than when there was no familiar pair. Furthermore, in the Discrete New sequences, the eighth and ninth response, which also formed a familiar pair, were only slightly faster than the real new pairs in the New sequences (i.e., R_2 and R_3, and R_5 and R_6). Notice that, despite the small magnitude of the effects, these data also offer support for the notion that motor chunks rely on associations between the responses that had been loaded together in the motor buffer during practice. Responses that had been separated by RSIs in the practice phase and, hence, had been part of different response groups, seem not to be associated.

Close examination of the individual response times in the Unstructured and Discrete six-key groups of ABB and AAB also demonstrates that there was little gain in using familiar parts in new contexts when compared with entirely new sequences. Once again, only a minor advantage was found of familiar parts in that the second response of familiar pairs and the second and third response of familiar triplets were faster. In addition, less errors were found in Discrete ABB and AAB sequences than in Discrete and New sequences. In sum, the minor net effect of familiar parts on sequence production lends support to Sternberg et al.'s (1990) contention that the occurrence of parts of an existing chunk does not substantially contribute to performance when embedded in otherwise new sequences.

Notice that the transfer phase also confirms results from the practice phase:

1. Within-group response times associated with Practiced response groups were longer in the Unstructured than in the Discrete condition and, again, this difference was more pronounced for three- than for six-key groups. This yields further evidence for the notion of concurrent preparation in the Unstructured condition, which has been discussed before in the practice phase.

2. The last response in the Practiced six-key group of the Discrete condition and, to a lesser degree, of the Unstructured condition, was relatively fast. This suggests concurrent retrieval of individual keypresses from the motor buffer (Verwey, 1994).

The analyses of correlations between individual response patterns in the practice phase and the transfer phase also yielded some interesting results. In the Unstructured 333 condition, the prediction was confirmed that participants who had grouped more in the Unstructured condition of the practice phase—and, hence, showed more clear indications for chunk development—had faster second responses in familiar pairs in New, ABB, and AAB than participants who had grouped less during practice. No such effects were found in the Discrete 333 condition, which only provides further support for the notion that when short response groups are prepared in advance, there is little effect of whether or not appropriate chunks exist.

However, in the 36 condition the usage of chunks in the Unstructured conditions of the practice phase appeared to have had little effect on the way in which familiar pairs were executed. The only significant correlation occurred in the R_8 R_9 pair of the New condition, but it was in the opposite direction; that is, the second response of this pair was slower, as the participant had grouped more during practice. The fact that this was observed in the Unstructured as well as in the Discrete condition does support the notion that execution of later elements in a new sequence are selected on-line, irrespective of whether the sequence could be prepared or not. A post hoc explanation is that participants who had less developed chunks still recognized the pair and tried to prepare it as a group (slow first and fast second response), whereas participants with better developed chunks could benefit from automatic priming of the second response (cf. MacKay, 1982) and, hence, did not attempt to prepare the pair as a whole.

Why would familiar pairs have had an effect in the 333 condition and little or no effect in the 36 condition? One explanation is that 333-condition participants could use some parsing mechanism, or rule (cf. Restle, 1970), which made parts of existing chunks available. This rule could be used effectively in the ABB and AAB conditions of the 333 group, as the structure of these sequences allowed consistent application of the rule across the new groups. In the 36 condition, however, little use could be made of existing chunks, as there was no single rule that could be used for extracting the various familiar parts that occurred in the new sequences.

In summary, the transfer phase offers further support that content-specific motor chunks had developed in the practice phase. Once again, motor chunks appear especially important when relatively long response groups are to be produced and when this occurs under time pressure. These data resolve the apparent contradiction between studies showing primarily content-specific effects of practice (e.g., Fischman & Lim, 1991; Logan, 1988) and those showing predominantly content-aspecific effects of practice (Chamberlin & Magill, 1992b; Verwey, 1992, 1994) by distinguishing between the development of a content-specific motor chunk (which can be loaded into the motor buffer in a single processing step) and content-aspecific effects of practice (attributed to more efficient advance buffer loading and the use of concurrent preparation and retrieval). That motor chunks are highly content specific is supported by the finding that when successive responses once belonged to different response groups, or when new response groups enclosed parts of familiar response groups, performance did

not improve much as compared with entirely new response groups. The fact that there was some improvement anyway may be due to the possibility of using some rule to extract the necessary bits and pieces from existing chunks.

General Discussion

This study examined effects of practice in generating movement sequences. It shows when the buffer loading and chunking concepts can be applied. Moreover, this study illustrates the relationship between buffer loading and motor chunking: Motor buffer models (e.g., Henry & Rogers, 1960; Sternberg et al., 1978) can predict performance of relatively short movement sequences that can be prepared in advance. The possibility of preparing short sequences in advance renders performance relatively insensitive to previous practice with the sequence at hand. In this situation, performance appears more dependent on aspecific effects of practice such as the capability of making efficient use of advance loading of the motor buffer and the possibility of retrieving individual elements from the motor buffer during execution of the preceding element (Semjen, 1992; Verwey, 1993, 1994, 1995).

When sequences are relatively long or when there is little time for advance preparation, content-specific effects of practice come into play. These effects are described by models stating that integrated representations, motor chunks, develop with consistent practice of sequential movement patterns (MacKay, 1982, 1987; Wickelgren, 1969). Such movement patterns are encountered in many real-world tasks such as speech and handwriting, and in the proficient control of various systems (e.g., vehicles and computers). These tasks are characterized by time pressure, which minimizes the opportunity for advance preparation of individual movements and increases the role of "off-the-shelf" motor chunks. Hence, in the absence of the appropriate motor chunks, the performer is loaded extensively by the need to rapidly select individual movements, whereas decision load for an expert is much lower, as he or she can choose one out of a large repertoire of content-specific chunks. Notice that the distinction between executing individual movements and motor chunks relates to the level of control: Whereas the novice is forced to control tasks at the level of individual movements, the expert can exert control at a higher level by selecting movement patterns rather than individual movements.

What Is a Chunk?

The term *motor chunk* suggests that other types of chunks may exist as well. Indeed, chunks have often been associated with facilitation of the processing and memorization of complex stimulus patterns (e.g., Miller, 1958; Reber, 1967). For example, chunking has been interpreted in terms of "a natural, maybe automatic, tendency to process stimuli by parts" (Servan-Schreiber & Anderson, 1990, p. 592). As yet, it is unclear to what extent motor chunks and chunks required for stimulus processing and memorization (i.e., non-

motor chunks) share properties. However, there are various indications that the grouping of knowledge representations into chunks occurs at various levels of information processing of which the motor level is but one (MacKay, 1982). This might explain the individual differences in the practice phase of our study in which some participants seemed to rely heavily on motor chunks, as indicated by high group-start to within-group ratios, whereas some others may have used a sequence representation at a different level that was not indicated by grouping of the sequence. Strategical differences between participants during practice may be an important determinant of the eventual form of chunks.

In his general model of behavior sequences, MacKay (1982) asserts that chunks exist at various levels of processing (e.g., at the phonological, syllable, and lexical levels) and are combined in hierarchical representations. Selection of a higher level chunk would automatically activate lower level chunks. This multilevel view on the acquisition of skill is in line with observations that skills can also be acquired by event observation (e.g., Fendrich, Healy, & Bourne, 1991; Howard, Mutter, & Howard, 1992) and mental practice (Decety & Ingvar, 1990; Feltz & Landers, 1983). It is also compatible with the notion that "individual movements that comprise the skill are first perfected to the point where they can be made more rapidly and accurately with little variation. Then they become welded together into 'chunks'" (Gallistel, 1980, p. 367): Only after individual movements can be performed rapidly because lower level chunks have developed, chunking at a higher level becomes feasible.

It is unclear what happens with low-level chunks when high-level chunks develop. One option is that lower level chunks can still be used in isolation. An alternative option is that consistent execution of a set of motor chunks in a fixed order yields larger chunks, which encompasses the individual elements of the original chunks, rather than that lower level chunks become part of a higher level chunk. This option might have the effect that the original parts are no longer available for use in isolation. An indication for the latter option has been reported by Zimmer and Körndle (1988). They demonstrated that transfer to the original part skills of riding a so-called pedalo (i.e., bicycle-like vehicle) decreased at higher skill levels and suggested that lower level chunks are no longer available once integrated into a higher level chunk. For now, the general inference is that chunking may occur at various levels and that one may change the level of control by selecting lower or higher level chunks. Future studies should address in more detail whether or not low-level chunks remain available after large chunks have developed.

The How and What Mechanism

Taken together, these findings suggest a model of sequence production involving two independent mechanisms that develop with practice, a content-specific *what* and a content-aspecific *how* mechanism. The *what* mechanism indicates what movements are to be executed and relies on the availability of a repertoire of motor chunks (Adams,

1984; Allport, 1980; MacKay, 1982, 1987). As such, the *what* mechanism consists of knowledge structures that seem closely related to what has been called *procedural knowledge* (e.g., Anderson, 1983; Stadler, 1989; Willingham, Nissen, & Bullemer, 1989). Little is known on the properties of these chunks and to what extent these properties depend on the level of abstraction. Possibly, partitioning of long sequences in early practice—due to, for example, the initial application of a set of rules (Jones, 1981; Restle, 1970), on temporal separation (Verwey & Dronkert, 1996; this study), or on expectations (Bartz, 1979)—determines chunk development in late practice. This content-specific *what* mechanism may also be related to implicit serial learning that is observed when some statistical structure is concealed in otherwise random keypressing sequences (e.g., Cohen, Ivry, & Keele, 1990; Stadler, 1989, 1992; Willingham et al., 1989).

The *how* mechanism determines in what way the processes are coordinated that are required for translating memory representations into actual movements. Because the *how* mechanism is responsible for translating information, rather than that it comprises knowledge representations, it is aspecific with respect to the movements produced. Hence, tasks that rely heavily on this mechanism (e.g., executing short movement sequences) require little practice before reaching asymptotic performance. The *how* mechanism involves the possibility of loading the motor buffer in advance and the development of concurrent processing (i.e., concurrent preparation and concurrent retrieval).

The *how* and *what* mechanisms are related in that carrying out a sequence of movements is most efficient when both can be used. However, in principle, they are independent and operate autonomously. Because the *what* mechanism entails the possibility of retrieving a knowledge structure that is subsequently loaded into the motor buffer, it will mainly affect sequence initiation times. The *how* mechanism, on the other hand, is responsible for processes responsible for loading the buffer as well as for retrieving and executing the individual elements from the buffer. This means that the *how* mechanism affects initiation as well as execution times.

There are various indications for a dissociation between a *what* and a *how* mechanism.

1. In one study, two- and four-key pressing sequences were practiced in response to two stimuli (Verwey, 1992). After reversing stimulus–response mapping, sequence initiation time increased considerably and clearly exceeded initiation time of entirely new sequences. In contrast, within-sequence response times were equal to those in the practiced sequences, including the fast last keypress indicative for concurrent retrieval of information pertaining to the individual elements. The increase in initiation time suggest that, in the Practice condition, the chunk representing the movement sequence was activated automatically by stimulus presentation, which inhibited response selection in the reversal condition. Automatic response triggering, given appropriate preparations, by the imperative stimulus has also been suggested by findings with single responses (Kramer, Strayer, & Buckley, 1990; Pashler & Baylis, 1991; Shiffrin

& Dumais, 1981) but Verwey's (1992) findings suggest that entire chunks can also be triggered with practice whereas execution is not affected.

2. Another indication for a *what* and a *how* mechanism is that Brown and Carr (1989) and Verwey (1994) found different learning rates for initiation and interkey response times. This is expected only when different mechanisms are involved in initiation and execution. Future work should more specifically address learning rates as a means of separating different mechanisms.

3. When comparing Unstructured response groups of the New sequences in the transfer phase (Figure 4) and unstructured response groups in Sessions 1 and 3 of the practice phase (Figure 2), these data show an advantage for the New response groups of the transfer phase in that group-start times were clearly slower in the practice phase. This is in line with the notion of a content-aspecific skill that developed in the practice phase and also improved performance in the New sequences of the transfer phase.

In short, this article offers evidence of the notion that two mechanisms develop when practicing fixed series of movements. The *how* mechanism is task aspecific and involves the possibility of processing information at various levels concurrently and loading the motor buffer in advance. These processes tap a single processing capacity. The *what* mechanism entails the development of content-specific motor chunks. Important questions for future research concern the merits of this model of sequence production in different tasks and the relationship to chunks at other levels of processing. Future research should also aim at further isolating mechanisms that are involved in the acquisition of skills and determining more precisely when these mechanisms really affect the learning process.

References

Adams, J. A. (1984). Learning of movement sequences. *Psychological Bulletin, 96,* 3–28.

Allport, D. A. (1980). Patterns and actions: Cognitive mechanisms are content-specific. In G. Claxton (Ed.), *Cognitive psychology: New directions* (pp. 26–64). London: Routledge & Kegan Paul.

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Bartz, A. E. (1979). Perceived task complexity and reaction time on a single trial and a series of trials. *Journal of Motor Behavior, 11,* 261–267.

Book, W. F. (1908). *The psychology of skill.* Missoula: Montana Press.

Brown, T. L., & Carr, T. H. (1989). Automaticity in skill acquisition: Mechanisms for reducing interference in concurrent performance. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 686–700.

Bruner, J. S. (1973). Organization of early skilled action. *Child Development, 44,* 1–11.

Chamberlin, C. J., & Magill, R. A. (1992a). A note on schema and exemplar approaches to motor skill representation in memory. *Journal of Motor Behavior, 24,* 221–224.

Chamberlin, C. J., & Magill, R. A. (1992b). The memory representation of motor skills: A test of schema theory. *Journal of Motor Behavior, 24,* 309–319.

Cohen, A., Ivry, R. I. & Keele, S. W. (1990). Attention and struc-

ture in sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 17–30.

Coover, J. E. (1923). A method of teaching typewriting based on a psychological analysis of expert typing. *National Education Association Addresses and Proceedings, 61,* 561–567.

Decety, J., & Ingvar, D. H. (1990). Brain structures participating in mental simulation of motor behavior: A neuropsychological interpretation. *Acta Psychologica, 73,* 15–34.

Drummond, H. (1981). The nature of and description of behavior patterns. In P. P. G. Bateson & P. H. Klopfer (Eds.), *Perspectives in ethology* (Vol. 4, pp. 1–33). New York: Plenum Press.

Estes, W. K. (1986). Array models for category learning. *Cognitive Psychology, 18,* 500–549.

Feltz, D. L., & Landers, D. M. (1983). The effects of mental practice on motor skill learning and performance: A meta-analysis. *Journal of Sport Psychology, 5,* 25–57.

Fendrich, D. W., Healy, A. F. & Bourne Jr., L. E. (1991). Long-term repetition effects for motoric and perceptual procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 137–151.

Fentress, J. C. (1983). Ethological models of hierarchy and patterning of species-specific behavior. In E. Satinoff & P Teitelbaum (Eds.), *Handbook of behavioral neurobiology* (Vol. 6, pp 185–234). New York: Plenum Press.

Fentress, J. C. (1984). The development of coordination. *Journal of Motor Behavior, 16,* 99–134.

Fischman, M. G., & Lim, C.-H. (1991). Influence of extended practice on programming time, movement time, and transfer in simple target-striking responses. *Journal of Motor Behavior, 23,* 39–50.

Frensch, P. A. (1994). Composition during serial learning: A serial position effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 423–442.

Fromkin, V. A. (Ed.). (1981). *Errors of linguistic performance: Slips of tongue, ear, pen, and hands.* New York: Academic Press.

Gallistel, C. R. (1980). *The organization of action: A new synthesis.* Hillsdale, NJ: Erlbaum.

Gee, J. P., & Grosjean, F. (1983). Performance structures: A psycholinguistic and linguistic appraisal. *Cognitive Psychology, 15,* 411–458.

Gordon, P. C. & Meyer, D. E. (1987). Control of serial order in rapidly spoken syllable sequences. *Journal of Memory and Language, 26,* 300–321.

Gottsdanker, R., Perkins, T., & Aftab, J. (1986). Studying reaction time with nonaging intervals: An effective procedure. *Behavior Research Methods, Instruments, & Computers, 18,* 287–292.

Henry, F. M., & Rogers, D. E. (1960). Increased response latency for complicated movements and a 'memory drum' theory of neuromotor reaction. *Research Quarterly, 31,* 448–458.

Howard, J. H., Mutter, S. A., & Howard, D. V. (1992). Serial pattern learning by event observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1029–1039.

Hulstijn, W., & Van Galen, G. P. (1983). Programming in handwriting: Reaction time and movement time as a function of sequence length. *Acta Psychologica, 54,* 23–49.

Hulstijn, W., & Van Galen, G. P. (1988). Levels of motor programming in writing familiar and unfamiliar symbols. In A. M. Colley & J. R. Beech (Eds.), *Cognition and action in skilled behaviour* (pp. 65–85). Amsterdam: North-Holland.

Jones, M. R. (1981). A tutuorial on some issues and methods in serial pattern research. *Perception & Psychophysics, 30,* 492–504.

Keele, S. W. (1986). Motor control. In K. R. Boff, L. Kaufman, &

J. P. Thomas (Eds.), *Handbook of perception and performance: Vol. II. Cognitive Processes and Performance* (pp. 30.1–30.60). New York: Wiley.

Keele, S. W., & Summers, J. J. (1976). The structure of motor programs. In G. E. Stelmach (Ed.), *Motor control: Issues and trends* (pp. 109–142). New York: Academic Press.

Kornblum S. (1965). Response competition and/or inhibition in two-choice reaction time. *Psychonomic Science, 2,* 55–56.

Kramer, A. F., Strayer, D. L., & Buckley, J. (1990). Development and transfer of automatic processing. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 505–522.

Lahy, J. M. (1924). Motion study in typewriting. In *Studies and reports: Series J (Educational), 3.* Geneva: International Labour Office.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). New York: Wiley.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 883–914.

MacKay, D. G. (1982). The problems of flexibility, fluency, and speed–accuracy trade-off in skilled behavior. *Psychological Review, 89,* 483–506.

MacKay, D. G. (1987). Asymmetries in the relationship between speech perception and production. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 301–333). Hillsdale, NJ: Erlbaum.

Machlis, L. (1977). An analysis of the temporal pattern of pecking in chicks. *Behaviour, 63,* 1–70.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Journal of Experimental Psychology, 56,* 485–491.

Miller, G. A. (1958). Free recall of redundant strings of letters. *Journal of Experimental Psychology, 56,* 485–491.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of Behavior.* New York: H. Holt.

Newell, A., & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1–55). Hillsdale, NJ: Erlbaum.

Paillard, J. (1960). Neurophysiology. In H. W. Magoun (Ed.), *Handbook of physiology* (Section 1, Chapter 67). Washington, DC: American Physiological Society.

Pashler, H., & Baylis, G. (1991). Procedural learning: 1. Locus on practice effects in speeded choice tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 20–32.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 5,* 855–863.

Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological Review, 77,* 481–495.

Rosenbaum, D. A., Kenney, S. B., & Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 86–102.

Sanders, A. F. (1990). Issues and trends in the debate on discrete vs. continuous processing of information. *Acta Psychologica, 74,* 1–45.

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

Schmidt, R. A. (1982). The schema concept. In J. A. S. Kelso (Ed.), *Human motor behavior: An introduction* (pp. 219–235). Hillsdale, NJ: Erlbaum.

Schneider, W. (1988). Micro Experimental Laboratory: An integrated system for IBM-PC compatibles. *Behavior Research Methods, Instrumentation, and Computers, 20,* 206–217.

Schneider, W., & Fisk, A. D. (1983). Attention theory and mechanisms for skilled performance. In R. A. Magill (Ed.), *Memory and control of action* (pp. 119–143). Amsterdam: North-Holland.

Segalowitz, S. J., & Graves, R. E. (1990). Suitability of the IBM XT, AT, and PS/2 keyboard, mouse, and game port as response devices in reaction time paradigms. *Behavior Research Methods, Instruments, & Computers, 22,* 283–289.

Semjen, A. (1992). Plan decoding and response timing during execution of movement sequences. *Acta Psychologica, 79,* 255–273.

Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 592–608.

Shiffrin, R. M., & Dumais, S. T. (1981). The development of automation. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 111–140). Hillsdale, NJ: Erlbaum.

Stadler, M. (1989). On learning complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1061–1069.

Stadler, M. A. (1992). Statistical structure and implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 318–327.

Sternberg, S., Knoll, R. L., & Turock, D. L. (1990). Hierarchical control in the execution of action sequences: Tests of two invariance properties. In M. Jeannerod (Ed.), *Attention and performance XIII* (pp. 3–55). Hillsdale, NJ: Erlbaum.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117–152). New York: Academic Press.

Teulings, H. L., Mullins, P. A., & Stelmach, G. E. (1986). The elementary units of programming in handwriting. In H. S. R. Kao, G. P. Van Galen, & R. Hoosain (Eds.), *Graphonomics: Contemporary research in handwriting* (pp. 21–32). Amsterdam: North-Holland.

Thomassen, A. J. W. M., & Van Galen, G. P. (1992). Handwriting as a motor task: Experimentation, modelling, and simulation. In J. J. Summers (Ed.), *Approaches to the study of motor control and learning* (pp. 113–144). Amsterdam: Elsevier.

Van Galen, G. P. (1991). Handwriting: Issues for a psychomotor theory. *Human Movement Science, 10,* 165–191.

Verwey, W. B. (1992). *Effects of practice in selecting and executing keypressing sequences* (Technical Rep. IZF 1992 B-10). Soesterberg, The Netherlands: TNO Institute for Perception.

Verwey, W. B. (1993). Effects of extended practice in a one-finger keypressing task. *Acta Psychologica, 84,* 179–197.

Verwey, W. B. (1994). Evidence for the development of concurrent processing in a sequential keypressing task. *Acta Psychologica, 85,* 245–262.

Verwey, W. B. (1995). A forthcoming keypress can be selected while earlier ones are executed. *Journal of Motor Behavior, 3,* 275–284.

Verwey, W. B., & Dronkert, Y. (1996). Practicing a structured continuous keypressing task: Motor chunking or rhythm consolidation? *Journal of Motor Behavior, 28,* 71–79.

Welford, A. T. (1968). *Fundamentals of skill.* London: Methuen.

Wickelgren, W. A. (1969). Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review, 76,* 1–15.

Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 1047–1060.

Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design.* New York: McGraw-Hill.

Wing, A. M. (1978). Response times in handwriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 153–172). New York: Academic Press.

Zimmer, A. C., & Körndle, H. (1988). A model for hierarchically ordered schemata in the control of skilled motor action. *Gestalt Theory, 10,* 85–102.