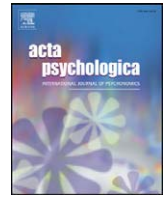




Contents lists available at ScienceDirect

Acta Psychologica

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

Chunking by colors: Assessing discrete learning in a continuous serial reaction-time task

Luis Jiménez^{a,*}, Amavia Méndez^a, Antoine Pasquali^{b,c}, Elger Abrahamse^d, Willem Verwey^d

^a University of Santiago, Spain

^b Free University of Brussels, Belgium

^c Adam Neurogenics, France

^d University of Twente, The Netherlands

ARTICLE INFO

Article history:

Received 19 November 2010

Received in revised form 25 March 2011

Accepted 27 March 2011

Available online xxx

PsychINFO classification:

2343 Learning & Memory

2346 Attention

Keywords:

Implicit learning

Sequence learning

Chunk learning

Statistical learning

ABSTRACT

Chunk learning (the process by which a sequence is learned and retrieved from memory in smaller, decomposed units of information) has been postulated as the main learning mechanism underlying sequence learning (Perruchet & Pacton, 2006). However, the evidence for chunk formation has been elusive in the continuous serial reaction-time task, whereas other continuous, statistical processes of learning account well for the results observed in this task. This article proposes a new index to capture segmentation in learning, based on the variance of responding to different parts of a sequence. We assess the validity of this measure by comparing performance in a control group with that of another group in which color codes were used to induce a uniform segmentation. Results showed that evidence of chunking was obtained when the color codes were consistently coupled to responses, but that chunking was not maintained after the colors were removed.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

When we learn a sequential skill, such as typing on a keyboard, reading a text or understanding a foreign language, we are usually not engaged in intentionally trying to encode the current episode for future reference, nor are we actively attempting to retrieve analogous experiences which could be brought to bear on the current situation. Rather, we typically just focus on the task, and try to do our best in responding to its actual demands. The fact that we can demonstrably improve our performance as we practice with such complex tasks, in the absence of clear intentions to learn or to rely on the retrieval of previous episodes, constitutes the mark of the phenomenon labeled as “implicit learning” (Reber, 1967).

One of the most enduring debates in the implicit learning literature concerns the role of attention in this form of learning. In a sequence learning task, for instance, participants respond to the locations of a series of stimuli by performing a corresponding series of key-presses. In order to respond, participants need to pay selective attention to respond to the locations of the stimuli, and in so doing they learn about the underlying sequences (Jiménez & Méndez, 1999). The fact that selective attention has been shown to be needed for implicit sequence

learning in most conditions (but see Cock, Berry, & Buchner, 2002; Perlman & Tzelgov, 2009, for contrasting cases), does not mean that this learning depends on the deployment of any other cognitive resources, beyond those required to perform the serial task. Thus, implicit learning would not depend on controlled processes such as rehearsal or hypothesis testing, but it would proceed automatically by binding together all aspects of successive experiences, precisely as they are shaped by selective attention.

1.1. Segmentation in sequence learning

Perruchet and his colleagues emphasized this point when they proposed what they called the “self-organizing consciousness” framework (e.g., Perruchet & Gallego, 1997; Perruchet & Pacton, 2006; Perruchet & Vinter, 2002). According to this framework, learning is driven by selectively attending to a reduced set of primitive units, which leads to the formation of a composite unit – or a *chunk* – by the concatenation of successively attended primitives. Once a chunk is created, it automatically acts as an attractor of attention, so that learners' attention is drawn toward the composite unit, which otherwise behaves like a primitive unit, thus allowing the composition of progressively larger chunks (Perruchet & Gallego, 1997). Importantly for this framework, these developed chunks are taken to be necessarily disjunctive (i.e., non-overlapping), due to the discrete nature of selective attention. Thus, even when participants

* Corresponding author at: Facultad de Psicología, Universidad de Santiago, 15782, Santiago de Compostela, Spain. Tel.: +34 98 156 3100.

E-mail address: luis.jimenez@usc.es (L. Jiménez).

are faced with a statistically homogeneous structure (i.e., a structure in which all its legal transitions are equally likely) learning would necessarily start by partitioning the sequence into a number of discrete segments, which could later be composed into progressively larger chunks.

These ideas were incorporated by Perruchet and Vinter (1998) into PARSER, a computational model first aimed at accounting for the statistical learning produced in artificial language learning tasks (Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996). In their studies, Saffran and her collaborators presented participants with a continuous stream of syllables containing no prosodic cues, but in which certain groups of trisyllabic strings were consistently repeated in the same order, thus forming what could be taken as “words” from an artificial language. As a result of training with this structure, the authors found that both children and adults were able to extract such hidden regularities, to the point of successfully discriminating these frequent “words” from other trisyllabic strings that appeared less often in the transitions between successive words. Perruchet and Vinter showed that PARSER could account for those learning effects by relying on the following elementary principles: (1) participants' attention is randomly focused on a small number of consecutive units at each time; (2) attention to a small group of units produces a new perceptual unit when the activation exceeds a certain threshold, (3) weight is gained each time a unit gets selected for attention, and (4) forgetting and interference cause losing weight for non-selected units, and for units corresponding to unselected segments embedded in a larger, selected unit. The simulations conducted by the authors showed that these attentional and associative principles accounted well for participants' ability to identify the most frequently repeated segments, since the units corresponding to such “words” yielded higher weights than those reached by the less frequent, “non-word” counterparts.

1.2. Chunking in DSP and SRT tasks

The question of whether sequence learning proceeds always in a similar way – that is, by decomposing sequences into segments – has been investigated in other sequence learning paradigms in which participants are required to respond to each stimulus, instead of being passively exposed to a stream of events. For instance, in both the Serial Reaction-Time (SRT) task (Nissen & Bullemer, 1987) and the Discrete Sequence Production (DSP) task (Verwey, 1999, 2001), participants are presented with a series of stimuli appearing on each trial at one of a reduced number of locations, and they are told to respond to these locations by pressing a series of spatially corresponding keys. In both cases, the series of stimuli follows a regular sequence, and reaction times (RT) reveal that participants learn about those sequences as a result of responding to each trial. Even though the requirement of responding to each trial results in a delay between successive trials which may hinder the joint activation of their representations in memory, this joint activation could still be produced with the aid of response codes (Ziessler & Nattkemper, 2001), thus resulting in chunks involving either single trials or pairs of trials which could later become concatenated with the following segment. The evidence for segmentation effects, however, is not equally clear in each of these two paradigms. While segmentation effects seem to arise spontaneously in the DSP task, they have not been observed as clearly in the SRT task.

In the DSP task, participants are required to respond to a reduced number of sequences of a limited length (for instance, two sequences of 3 and 6 elements) which are repeated extensively, separated by pauses, and sometimes marked by sequence-specific cues (Verwey, 1999). Responses to each trial within a given sequence are also marked by key-specific cues, and the response-to-stimulus interval (RSI) is usually set to 0, thus encouraging the internal grouping of the set of responses contained within a sequence. In these conditions,

which tend to favor an explicit encoding of the sequence, and an internal (probably, working-memory based) control of performance, segmentation arises even when no pauses are interposed between successive sequences. The results obtained with the DSP task suggests that the amount of elements represented within a motor chunk is limited (e.g., Bo & Seidler, 2009; Kennerley, Sakai, & Rushworth, 2004; Verwey, 2001, 2003a; Verwey, Abrahamse, & Jiménez, 2009). Moreover, Verwey (2010), (see also Verwey, Abrahamse, Ruitenberg, Jiménez, & de Kleine, in press) argued that elderly no longer used chunks in the DSP task whereas statistical learning seemed to continue. If these data appear reliable, this indicates that development and use of chunks and statistical learning are independent sequence learning processes.

In contrast, evidence for segmentation effects has been scarcer in the SRT task. In fact, it has only been documented when pauses were introduced to separate different segments (Stadler, 1993, 1995), or when relational properties were arranged to increase the salience of the segments (Koch & Hoffmann, 2000). Moreover, the evidence for chunk learning in this latter case is not very compelling, since the purported segmentation effects have been confounded with preexistent response tendencies (Jiménez, 2008, but see Kirsch, Sebald, & Hoffmann, 2010). In any case, given that including such relational properties improves explicit recognition of the sequences (and therewith working-memory control on performance), it is still a matter of debate whether implicit sequence learning as it arises in a typical SRT task should be understood either as the result of a discrete process of chunking (perhaps with individual differences concealing chunking patterns; cf. Verwey, 2003b; Verwey & Eikelboom, 2003), or as the effect of a continuous process of statistical learning.

1.3. Continuous statistical learning vs. discrete chunk learning

The evidence in favor of a natural tendency to break a long sequence into a series of discontinuous segments should be tested against alternative proposals suggesting that sequence learning arises continuously out of learners' sensitivity to the conditional probabilities between successive events (i.e., statistical learning). For instance, connectionist models such as the Simple Recurrent Network (SRN, see Elman, 1990; Cleeremans, 1993) assume that responding to a structured material produces a continuous process of learning about the conditional probabilities of each successor in each possible context. The SRN exploits the back-propagation algorithm to learn continuously to predict the successor of a series, based on the current input plus a distributed representation of its context, which is instantiated in terms of a set of recurrent connections within a hidden layer. Christiansen, Allen, and Seidenberg (1998) showed that such a model can account for the statistical learning reported by Saffran, Aslin, et al. (1996), Saffran, Newport, et al. (1996), and many other studies have shown that this model fits well with the results of a variety of paradigms within the SRT task (Cleeremans, 1993; Cleeremans & Jiménez, 1998; Jiménez, Méndez, & Cleeremans, 1996). Thus, as highlighted by Perruchet and Pacton (2006), the available evidence up to now is compatible at least with three alternative scenarios concerning the nature of sequence learning: (1) chunk learning and statistical learning are independent mechanisms, with the former one corresponding more clearly to the results of an intentional search, whereas the latter one accounts better for the results obtained in more implicit tasks; (2) statistical learning and chunk learning are two successive steps, with chunk boundaries arising out of statistical computations; or (3) learning always proceeds by means of a chunking process, and the observed sensitivity to the statistical properties of the material arises as a by-product of chunk learning.

1.4. Disentangling the predictions from chunk vs. statistical learning

Unfortunately, disentangling the predictions arising from these three scenarios has proved to be a difficult endeavor. First, both a continuous mechanism of statistical learning and a discrete process of chunking may

yield equivalent representations of the underlying sequence by the end of learning, if the chunks are allowed to enter into successive processes of composition of larger chunks (Perruchet & Gallego, 1997). Moreover, if both models are exposed to a sequential structure containing transitions of different conditional probabilities, such as those employed in Saffran's studies, the predictions arising from both models also tend to converge at the acquisition stage, for both of them predict that the more reliable transitions would be acquired faster and more strongly than the comparatively less reliable transitions. Thus, we contend that the predictions derived from these two types of models are easier to disentangle when they face the acquisition of a homogeneous sequence, in which all its legal transitions are equally likely.

Consider, for instance, a Second-Order Conditional (SOC) sequence composed of 12 trials defined over six different locations, such as that represented by the digits 2-5-4-6-1-3-5-2-4-1-6-3 (see Reed & Johnson, 1994, for an extensive analysis of SOC structures). In this sequence, all individual items are equally frequent, and all legal transitions are equally predictable by relying on a given number of previous trials. Thus, if we consider contexts made up of a single trial, two possible successors become equally likely at this point (e.g., locations "3" or "6" are both equally likely after location "1"). If, on the other hand, we consider a context made up of two or more previous trials (e.g., the sub-series "6-1") this reduces the options to a single legal successor (which, in this case, is location 3). In such a sequence, therefore, and assuming the absence of systematic differences in salience between parts of the sequence, a continuous process of statistical learning would tend to produce a rather homogeneous improvement of performance over the whole sequence, thus mirroring the homogeneous statistical structure of the sequence. In contrast, a chunk-based learning model such as PARSER would still predict that participants' attention would get attracted to some particular segments of the series, and hence that sequence learning would start discontinuously from those selected segments. Only with more extensive training could the model predict that such disjunctive segments would tend to get integrated into progressively larger chunks, as a result of a composition process.

1.5. The present study

According to these arguments, we surmise that looking for discontinuities over the acquisition of knowledge about a homogeneous sequence seems like a good strategy to find evidence for the kind of discrete chunk learning process proposed by e.g. Perruchet and Gallego (1997); Perruchet and Pacton (2006); and Perruchet and Vinter (1998, 2002). Importantly, however, if segments are created more or less arbitrarily across learners (depending only on random or idiosyncratic factors) then their segmentation effects would tend to cancel out among a group of learners, thus rendering their effects undetectable on the average measures of RTs. In an attempt to solve this problem, we developed a measure of segmentation inspired by the rationale from the Analysis of Variance, which may be sensitive to chunk learning even if each participant breaks the sequence in a different way. To examine the validity of this measure we used it not only in the typical conditions of the SRT task (which may or may not produce segmentation effects), but also under conditions in which a common segmentation pattern was promoted on all participants. This allowed us to assess chunk learning simultaneously with measures of central tendency and with the alternative measure proposed here.

As for the alternative index, we surmise that a measure based on the variance of a distribution of responses might be better suited to capture segmentation effects in conditions in which each participant could be breaking the sequence into different chunks. Indeed, if different participants tend to learn selectively about distinct segments of the sequence, such discontinuous learning effects would tend to disappear when the learning scores are averaged over learners. Yet, discontinuous performance could be observed through the measures of variance, which would tend to increase if people start learning selectively about certain segments. The variance computed for each participant will

steadily increase as the number of segments approximates to half of the number of total sequence elements, thus leading to a bimodal distribution in which roughly half of the elements would be responded to faster (because they would be internal to a chunk) whereas the other half would be responded to more slowly, as they would correspond to transitions between successive chunks.

Moreover, given that the variance can increase or decrease over training because of many other factors, we propose to further isolate the amount of variance in RT attributable to chunk learning by comparing two components: the variance concerned with differences in responding to distinct parts of the sequence, and the residual variance caused by any other factor, such as practice or fatigue, which may affect repeated responses to the same element of the sequence. According to this distinction, we consider the ratio between these two components of variance as a new score that is potentially useful to reflect chunk-based performance. Thus, over a given block containing n repetitions of a sequence of m elements, we define the *Variance between elements* as the variance in RT due to responding differently to distinct elements of the sequence. More specifically, this component is taken as the variance, computed over the m elements of the sequence, of the mean RTs calculated over the n repetitions of that element over a given block,

$$Variance_{between} = \frac{\sum_{j=1}^m \left(\frac{\sum_{i=1}^n x_{ji}}{n} - \bar{X} \right)^2}{m-1}.$$

In contrast, the *Variance within elements*, which serves as a control for other sources of variance, is taken as the mean, computed over the m elements of the sequence, of the variances calculated for the n repetitions of each element over a given block:

$$Variance_{within} = \frac{\sum_{j=1}^m \frac{\sum_{i=1}^n (x_{ji} - \bar{X}_j)^2}{n-1}}{m}.$$

It can be demonstrated that the sum of these two components of variance amounts to the overall variance of performance when the variance scores are computed as population parameters (i.e., dividing by the number of observations, instead of by the number of observations minus one). The sum of these two indices does only approximate to the overall variance when the scores are computed as sample statistics, as we do here, but the equation could be fitted with a simple scaling operation (see Appendix A for a formal proof):

$$Variance_{total} = \frac{n \cdot (m-1)}{n \cdot m - 1} Variance_{between} + \frac{(n-1) \cdot m}{n \cdot m - 1} Variance_{within}.$$

We surmise that the ratio between these two scores can therefore be taken as a sensible index of how much of the variance in participants' performance can be attributed to differences in the way in which they respond to distinct elements of the sequence, as compared to the variance attributable to other, unspecific causes which might affect responses over training.

As for the manipulation used to force a common segmentation pattern over a group of participants, we did not use pauses, as generally arranged in the DSP task, because pauses could be prone to rhythm learning or inhibition effects, and they might interfere with the associative processes responsible for the accrual of statistical information. Instead, we started from Koch and Hoffmann's (2000) idea of increasing the salience of certain segments without compromising the homogeneous statistical structure of the series. However, instead of relying on recurrent patterns such as repetitions, reversals, or ascending and descending runs, which have been found to produce RT differences independent of learning (Jiménez, 2008), we relied on an external source of communality, based on an additional

Control group, in which the color cues fell on different segments for each repetition. If this was confirmed, one could expect to obtain slower RT in response to the first trial of each triplet, not just because those trials required a switch in target color (note that color transitions were also predictable), but because color switches corresponded to transitions between chunks. To assess this effect independently from the switches in color cues, we included a transfer phase in which all colors were removed, and participants were told to respond to a black square following the same trained sequence, while the remaining squares were filled with the background color.

One could think of the group manipulation as producing two alternative effects. First, if learning proceeds always as a chunking process, according to the third scenario described by Perruchet and Pacton (2006), then providing participants with color cues systematically associated with the sequence of responses would not necessarily tend to make learning more fragmentary than in the Control condition, but it would only make the segment boundaries more uniform across learners. If that were the case, then the measure of variance which assesses the relative dispersion of RT attributable to learning discrete segments would show comparable levels in both groups, but the measures of average RT in response to each element of the sequence would tend to yield more evidence of chunk learning in the Uniform Segmentation group.

In contrast, if statistical learning arises as a default implicit learning process in both groups, one could expect such a continuous learning to arise in both conditions, perhaps supplemented with more explicit chunk learning in the Uniform Segmentation group, in which a systematic relation was arranged between each color cue and a specific triplet. In that case, if the salient association between colors and responses were responsible for the discrete learning effects, then an increase in the variance of performance will be expected to arise in this condition. To analyze the contents of the resulting knowledge, we included test blocks at specific points within the training and transfer phases, in which participants responded to segments generated either in accordance or against the constraints of the sequence. At the end of the experiment, explicit knowledge was tested by means of direct measures of sequence learning, which were taken from segment recognition and generation tasks.

2. Method

Participants performed a color-matching variant of the SRT task (Rünger & Frensch, 2008). On each trial, a colored square appeared centered above a horizontal row containing a group of six colored squares. Participants had to find, on the horizontal row, the specific square matching the color cue presented at the top, and to respond by pressing the key corresponding to its location in the bottom row (see Fig. 1). The color cue could be red, green, blue, or yellow, and these cues were presented in a fixed order so that they were predictable by themselves. The horizontal row contained six squares with colors that varied on every trial. The colors included the four possible targets plus two distractor colors (pink and orange). Unknown to the participants, the series of responses required over successive trials (i.e., the locations of the target colors on the bottom row) followed a deterministic SOC sequence of 12-elements over the training blocks. For participants in the Uniform Segmentation group, the color cues were systematically associated to the same triplets within the SOC sequence. Thus, the color cues involved four successive triplets of red, green, blue, and yellow squares, associated respectively with the first, second, third and fourth triplet of responses/locations. The Control group was presented with a similar succession of color cues, but in this case the red cue was presented for only two trials, thus resulting in a series of 11 color cues which run out of phase, and thus uncorrelated, with the sequence of 12 responses/locations.

At blocks 1, 6, and 11, test blocks were included to assess both sequence learning and chunking effects (see description below). After the training phase (i.e. blocks 1–8), all participants were presented with five transfer blocks (blocks 9–13) in which both the upper color

cues and the colors in the bottom row were removed. This left only a row of six white squares, of which one was filled with black on every trial. This allowed the sequence to be carried out as a standard SRT task. Test block 11 was used to assess the learning expressed under these transfer conditions. Finally, after the transfer phase, two generation blocks and a recognition block were administered to assess explicit knowledge of the sequence fragments.

2.1. Participants

Thirty six students from the University of Santiago participated in the study in exchange for a small monetary fee. To maintain their motivation to cope with task demands, they were informed that 10% of the participants who yielded the best scores would receive a doubled fee.

2.2. Procedure

Eighteen students were assigned to the Uniform Segmentation group, and the same number of participants was tested under the Control conditions. After preliminary instructions, participants started with the color task over 14 completely random trials (block 0), and then they performed a first test block (block 1, see upper table within Fig. 1) to confirm that responding to different types of trials was not different before exposition to the sequence. After this first test block, participants completed four training blocks (blocks 2 to 5) with the color arrangement assigned to their condition. A second test block (block 6) was presented at this point, so as to analyze the learning acquired over the color-cue task. This was followed by two more training blocks (blocks 7 and 8), and then participants were presented with the transfer phase.

At this point, participants were informed that the color would no longer be presented, but that they should now respond to a unique black square that appeared among white squares. Participants were presented with two of these transfer blocks (blocks 9 and 10) in which the training sequence of responses was maintained, followed by another test block (block 11) in which the sequence was altered, and two more transfer blocks (blocks 12 and 13) in which the sequence was reintroduced to prepare participants for the direct measures of learning. Two generation tasks (inclusion followed by exclusion, see Destrebecqz & Cleeremans, 2001) and a recognition task were then used to assess the extent to which participants could use their sequence knowledge when they were explicitly required to do so. On the inclusion task participants were told to generate a sequence resembling that which was more often seen over training. On the exclusion task they were told to try to generate a structure *opposite* to that seen over training, so that, if they remember to have seen the series “123”, they should try to generate precisely the inverse series, “321”. The inversion logic adopted in this variant of the exclusion task was preferred to the standard “avoiding” instructions, because it makes both tasks more comparable than the original inclusion and exclusion tasks. Indeed, the standard exclusion task has always been problematic for the PDP procedure because, in addition to the explicit “avoiding” instructions, it includes many other more subtle constraints, such as avoiding continuous repetitions, sampling all the keys with about the same frequency, or avoiding systematic ascending or descending runs, which are often considered but remain unsaid (i.e., Wilkinson & Shanks, 2004). In the present experiment, we aimed to put these constraints more explicitly, and so we arranged SOC sequences which had no SOC transition in common with their reverses, and asked participants to generate a sequence opposite to that experienced over training. In so doing, we were explicitly asking them to generate a structure which should be analogous to that found over training, but in which all SOC transitions were different than those seen over the training phase. Finally, on the recognition task we presented participants with several trial triplets, and required them to judge, using a scale from “1” to “6”, whether each of these triplets corresponded to triplets which they have seen repeatedly over the previous training.

On each training block there were 12 repetitions of the training sequence, for a total of 144 trials per block. The sequence was selected so that it contained no triple runs (i.e., ascending or descending series of three consecutive locations); no immediate repetitions of the same response, and no reversals (i.e., repetitions of the location presented just before the last trial, like 121). Moreover, it was built so that the four triplets associated with individual colors in the Uniform Segmentation group could be reordered and broken to generate the appropriate test blocks without incurring in repetitions, reversals, or triple runs.

The original sequence used to generate these trials was 254-613-524-163. To control for the specific difficulty of individual transitions and the fingers used, we transformed this original sequence into five additional sequences, by moving each location one step to the right (i.e., location 1 becomes 2 in the next sequence, and so on). Each one of the six resulting sequences was used as the training sequence for three out of the 18 participants in each condition. See Fig. 1 for details.

Over the test blocks, participants in both groups were presented with the same stimulus arrangement, designed to include trials in accordance with the learned chunks, together with other trials which broke these chunks. In the test blocks involving colors (i.e., blocks 1 and 6), all participants were presented with the same series of consecutive color cues used over the training blocks for the Common Segmentation condition (i.e., a series of three red, three green, three blue, three yellow cues) but the responses/locations were arranged as 8 series of 12 trials, including different types of transitions. Specifically, each test block included two series of 12 trials reproducing the training sequence (TS), two series of a sequence that we label as “Disordered Chunks” (DC), two repetitions of a sequence produced by breaking the chunks in their first transition (B1) and two more exemplars of a series produced by breaking the chunks in their second transition (B2). To control for linear practice or fatigue effects over the test series, we arranged the two repetitions of each series according to an ABBA design. Thus, the 8 series making up each test block were arranged as TS, DC, B1, B2, B2, B1, DC, and TS. The “Disordered Chunk” series were produced by altering the order of the triplets within the 12-trial sequence, while maintaining the integrity of the triplets. This resulted in a change of the association between each of these triplets and the color cues with which they were normally associated. The B1 series were built by replacing the first element of each triplet with that of some other triplet, so that this first transition was broken without creating repetitions, reversals or runs (e.g., “254” becomes “154” for participants 1, 7 and 13). Likewise, the B2 series were built by changing the third element of each triplet without incurring repetitions, reversals, or runs (e.g., “254” was replaced by “253”, see Fig. 1 for an illustration of all these changes).

After training with the appropriate color conditions, participants were shifted to the transfer phase without colors, and they were also presented with a test block without colors, which otherwise reproduced the structure described above. At the end of this transfer phase, participants were presented with the explicit measures of

sequence learning. Both the generation and recognition tasks were conducted with the transfer version of the task, that is, without colors. First, they were presented with a continuous generation task under inclusion conditions, in which they were told to generate 60 trials which followed a sequence as similar as possible to that which was most often presented over the training period. Next, they were presented with exclusion conditions, in which they were instructed to generate a series of 60 trials which were opposite to that experienced over the training trials. Finally, they were presented with a recognition task in which they first responded to, and then rated a set of 20 triplets presented in a random order, in terms of the degree in which they believed that each triplet had been (1) or had not been (6) part of the trained sequence. The 20 triplets corresponded, respectively, to the four chunks induced for that participant in the Uniform Segmentation group (Sequential chunks), the remaining 8 sequential triplets which corresponded to sequential transitions between chunks (Sequential transitions), and the 8 non-legal triplets which were used to produce the B1 and B2 illegal segments over the test blocks (see the table at the bottom of Fig. 1 for an example of the triplets presented to the participants trained with the first sequence).

3. Results

Results from the random block 0 were removed from all analyses, and so were the responses to the first two trials from each block, which were not predictable. Percentage of accurate responses was generally high over the remaining 13 blocks, averaging .97 in both groups. A mixed-design Analysis of Variance (ANOVA) conducted over the percentage of correct responses with Condition (2) as a grouping factor, and with Block (13) as a repeated measure showed no significant effect of condition, $F < 1$. There was a trend to improve accuracy over training, but this was not significant after applying the Greenhouse–Geisser correction, $F(12, 408) = 2.15$, $\eta_p^2 = .06$, $p = .09$.

As there was no evidence of a trade-off between reaction time and accuracy measures, we relied exclusively on the measures of reaction time (RT) on correct responses as the indirect measure of learning over both training and transfer phases. We will present separately the analyses of performance over training and transfer phases, the analyses of RT produced in response to either internal or external parts of a segment, and the effects of sequence learning observed over the test blocks. Then we will assess learning as expressed through generation and recognition tasks, and we will present the evidence for chunk learning provided by the measures of variance.

3.1. Performance over training and transfer phases

Fig. 2 shows that performance in both groups was initially equivalent over the initial test block, $F < 1$. An ANOVA restricted to the training

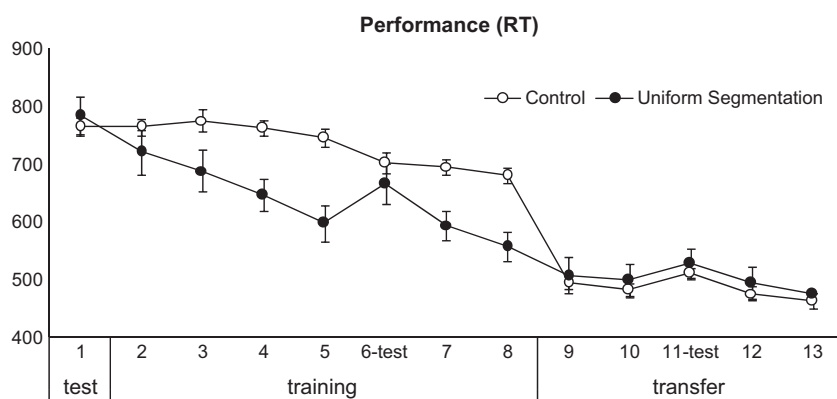


Fig. 2. Performance over training and transfer phases. Average RT for participants in the Control and Uniform Segmentation conditions, represented separately for successive training and transfer blocks, including the Test blocks arranged as Blocks 1, 6, and 11. Error bars represent Mean Standard Errors.

blocks before transfer (i.e., removing the test blocks 1 and 6), showed significant effects of Condition, $F(1, 34) = 12.50$, $\eta_p^2 = .27$, $p < .001$, and Block, $F(5, 170) = 48.80$, $\eta_p^2 = .59$, $p < .001$, as well as a significant Condition \times Block interaction, $F(5, 170) = 6.54$, $\eta_p^2 = .16$, $p < .01$. In contrast, a similar ANOVA conducted over the transfer blocks (also after removing the test block 11) showed no significant effect of Condition, $F < 1$. The effect of Block was also significant over the transfer phase, $F(3, 102) = 14.06$, $\eta_p^2 = .29$, $p < .001$, but there was no hint of a significant Condition \times Block interaction, $F < 1$. Plainly then, performing the SRT task with color cues associated with the response sequence improved performance over the training phase, but it did not produce enduring effects over the transfer phase after the colors were removed.

3.2. Segmentation in performance over training and transfer phases

To assess whether performing the SRT task with color cues associated with the response sequence facilitated its segmentation into triplets, and whether this segmentation survived the removal of the colors over the transfer phase, we analyzed performance in terms of the position of each trial within a color-cue triplet, as they were presented in the Uniform Segmentation group. This classification did not make sense for participants in the Control group, because they had been trained with a different distribution of color cues, and were presented with color-cue triplets only over the test blocks. As it is evident from an inspection of Fig. 3, the arrangement of color cues in triplets did dramatically improve responding to the second and third items from each triplet, and it did so, although to a lesser extent, when the regular series of color triplets were presented to the Control participants over the test block 6. In contrast, responding to the first element of a triplet in the Uniform Segmentation group produced a slow-down in response as compared to the RT produced in response to the same elements over the control group. This slowdown of responses to the first element of a triplet appeared to be maintained, at least partially, over the transfer phase, despite the removal of the colors.

Analyses conducted separately for the training and transfer phases after removing the test blocks, using Condition as a grouping factor, and Block and Trial (1st to 3rd) as within-participant variables showed that the three-way interaction was significant over training, $F(10, 340) = 6.93$, $\eta_p^2 = .17$, $p < .001$, but not over the transfer phase, $F < 1$. The overall effect of Condition was also significant over training $F(1, 34) = 12.36$, $\eta_p^2 = .27$, $p < .001$, but not over the transfer phase, $F < 1$. ANOVAs conducted separately for each condition over the training phase showed that, for the Uniform Segmentation group, there were significant effects of Trial $F(2, 34) = 46.53$, $\eta_p^2 = .73$, $p < .0001$, and Block, $F(5, 85) = 29.80$,

$\eta_p^2 = .64$, $p > .001$, as well as a significant Trial \times Block interaction, $F(10, 170) = 8.15$, $\eta_p^2 = .32$, $p > .001$. Helmert contrasts for the Trial effect confirmed that responding to the second and third elements from each triplet did not differ from each other in that group, $F < 1$, but that they were faster than responding to the first element $F(1, 17) = 52.84$, $\eta_p^2 = .78$, $p > .001$. In contrast, for the Control group no effect or interaction involving Trial approached significance, $F < 1$, and there was only a significant effect of Block, $F(5, 85) = 23.43$, $\eta_p^2 = .58$, $p > .001$.

As for the Transfer phase, the overall analysis showed a significant Trial \times Condition interaction, $F(2, 68) = 4.78$, $\eta_p^2 = .12$, $p < .05$. As can be observed from Fig. 3, the interaction stems mainly from the existence of a larger difference between responding to the first trial and responding to the second and third trials in the Uniform Segmentation group, as compared to that observed in the Control group. An ANOVA conducted on these differential scores (i.e., RT to the first trials minus the average of RT to the second and third trials), with Condition as a grouping factor and Transfer blocks as a within-participants variable, confirmed that these scores were larger for the Uniform Segmentation group than for the Control group, $F(1, 34) = 8.48$, $\eta_p^2 = .20$, $p < .01$. Thus, although responding to the first element of a triplet appears to have been faster than responding to the second and third trials in both groups, such a difference was larger in the Uniform Segmentation group.

3.3. Performance over the test blocks

Performance over the test blocks was used to assess both sequence learning in general, and the specific contents of this learning. As an overall measure of sequence learning expressed independently of the color cues, we focused on the learning manifested over the transfer phase, comparing performance over the test block 11 with that obtained over its two neighboring sequence blocks (10 and 12). The corresponding ANOVA with Condition and Type of Block (test vs. sequence) showed neither a significant effect of Condition, nor a significant interaction Condition \times Type of Block, $F < 1$. However, it showed a significant effect of Type of Block, $F(1, 34) = 28.01$, $\eta_p^2 = .45$, $p < .001$. This suggests that both groups obtained similar benefits from responding to the training sequence, as compared to responding to the mix of trials arranged in the test block 11.

As for the comparison between the different types of trials included within the test blocks, we computed the average RT in each of these types of trials (TS, DC, B1, B2), both over the transfer phase (with no colors) and over the learning phase (with color cues). The results are depicted in Fig. 4.

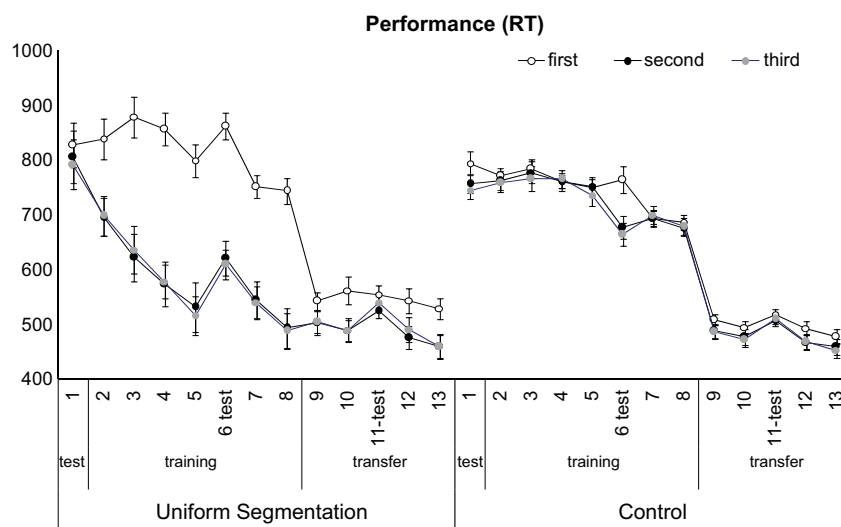


Fig. 3. Segmentation in Performance over training and transfer phases. Average RT for participants in the Uniform Segmentation and Control conditions, represented separately for successive training and transfer blocks, and for the first, second, and third trial from each triplet. Error bars represent Mean Standard Errors.

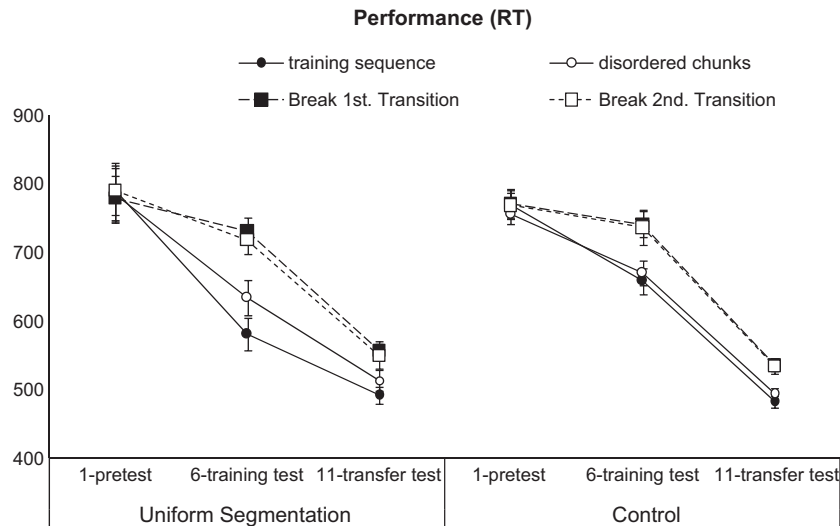


Fig. 4. Performance over the test blocks. Average RT for participants in the Uniform Segmentation and Control conditions, represented separately for the different types of trials included over the test blocks, and separately for each of the test blocks included before training, within the training phase, and within the transfer phase. Error bars represent Mean Standard Errors.

To exclude potential differences before training, a first ANOVA conducted over the Pretest block (1) showed no initial differences between Conditions or among types of trials, as well as no significant interaction between them ($F_s < 1$). A similar ANOVA conducted over the test block (6) showed a significant effect of Type of Trial, $F(3, 102) = 52.89$, $\eta_p^2 = .61$, $p < .001$, and a significant Condition \times Type of Trial interaction, $F(3, 102) = 3.71$, $\eta_p^2 = .10$, $p < .05$. As can be observed from Fig. 4, the interaction arises from the fact that participants in the Uniform Segmentation group responded faster to the training sequence trials (in which the triplets were still associated with the same color cues that appeared over training) than to the disordered chunks, in which the color-to-triplet association was not maintained. Consistently, separate ANOVAs conducted for each group over the test block 6 showed effects of Type of Trial in both the Uniform Segmentation group, $F(3, 51) = 39.33$, $\eta_p^2 = .70$, $p < .001$, and in the Control group, $F(3, 51) = 16.06$, $\eta_p^2 = .49$, $p < .001$, but the contrast between TS and DC trials was significant only in the Uniform Segmentation group, $F(1, 17) = 17.77$, $\eta_p^2 = .51$, $p < .01$. Finally, an ANOVA conducted over the transfer test block (11), without the colors, showed a significant effect of Type of Trial, $F(3, 102) = 45.22$, $\eta_p^2 = .57$, $p < .001$, but not any significant effect or interaction involving Condition ($F_s < 1$).¹

¹ It is noteworthy that, in the Control group, responding to the DS trials was faster than responding to the B1/B2 series, even though participants in that group had not been presented with any specific segmentation pattern. This could be attributed to the fact that by disordering the chunks we only introduced four changes in the second order conditionals, whereas by breaking the chunks we produced a larger number of changes in the statistical structure, both in the second order and in the first order conditionals. Alternatively, one might argue that these effects could reflect some kind of carry-over effect between successive trials, derived from the fixed order of presentation of the different types of trials within the test blocks. Indeed, an analysis comparing responses to the first vs. the second appearance of these series over the test block 6 indicated that the time of appearance affected performance, since there was a significant half \times type of trial interaction, $F(3, 102) = 34.98$, $\eta_p^2 = .51$, $p < .001$. Particularly, the difference between DS and B1/B2 trials was highly significant over the first half of the test block, $F(1, 34) = 117.58$, $\eta_p^2 = .78$, $p < .001$ but fell below significance over the second half, $F(1, 34) = 3.56$, $\eta_p^2 = .10$, $p = .07$. Notably, this decay of the relevant difference between halves was observed specifically in the Uniform Segmentation group, $F(1, 17) = 39.56$, $\eta_p^2 = .70$, $p < .001$, but did not reach significance in the Control group $F(1, 17) = 4.16$, $\eta_p^2 = .20$, $p = .06$. We attribute the carry-over effect observed specifically in the Uniform Segmentation group to the fact that participants' reliance upon their chunk information over the first presentation of the DS series tended to be abandoned after being exposed to a long series of non-structured trials latter in that block. In contrast, the difference observed in the Control group between DS and B1/B2, which was roughly maintained between halves, could be better interpreted as arising from the statistical differences that exist between these two types of sequences.

In sum, the analyses suggest that participants in the Uniform Segmentation group were sensible to the associations arranged between specific colors and response triplets, so that they responded better when the chunks were associated to the same colors. However, the results also indicate that training under this correlated condition did not produce a pervasive effect in shaping the particular segments of the sequence that participants learned, since the differences between groups observed over training were no longer observed over the transfer block, when the colors were removed.

3.4. Direct measures of learning

Both generation and recognition tasks were included at the end of the experiment to test whether participants in the Uniform Segmentation group were more able to generate or recognize those chunks which were rendered more salient by associating them with fixed colors, as compared to other sequential triplets encompassing transitions between chunks, or to the non-sequential triplets which had been seen exclusively over the test blocks (See Fig. 1, bottom Table for a collection of the relevant triplets as used for one specific sequence). Fig. 5 (upper panel) presents the average generation scores separately for inclusion and exclusion tasks, for participants in the Control and Uniform Segmentation groups, and for different types of triplets (chunks, other sequential triplets not trained as a chunk, and non sequential triplets used in B1 and B2 sequences). Because the number of possible segments within each category is not balanced (there were only four chunks, but 8 segments containing other sequential transitions, and 8 non-sequential test triplets), the number of generated segments produced over the 60 trials was normalized dividing by the number of possible cases of each type. Fig. 5 (bottom panel) represents the average recognition scores given to the same segments separately by participants in the Control and Uniform Segmentation groups. The axis is inverted in this case to indicate that low values (close to 1) represent higher recognition scores.

As for the Generation tasks, an ANOVA with Condition as a grouping factor, and with Task (inclusion vs. exclusion) and Type of triplet as repeated measures showed significant effects of Task, $F(1, 34) = 13.23$, $\eta_p^2 = .28$, $p < .001$, and Type of triplet, $F(2, 68) = 9.02$, $\eta_p^2 = .21$, $p < .001$, as well as a significant interaction Task \times Type of triplet, $F(2, 68) = 6.65$, $\eta_p^2 = .16$, $p < .01$, but no effect or interaction involving Condition ($F_s < 1$). Separate analyses for inclusion and exclusion tasks showed that the exclusion task did not discriminate between groups or triplets. In the inclusion task the effect of Condition

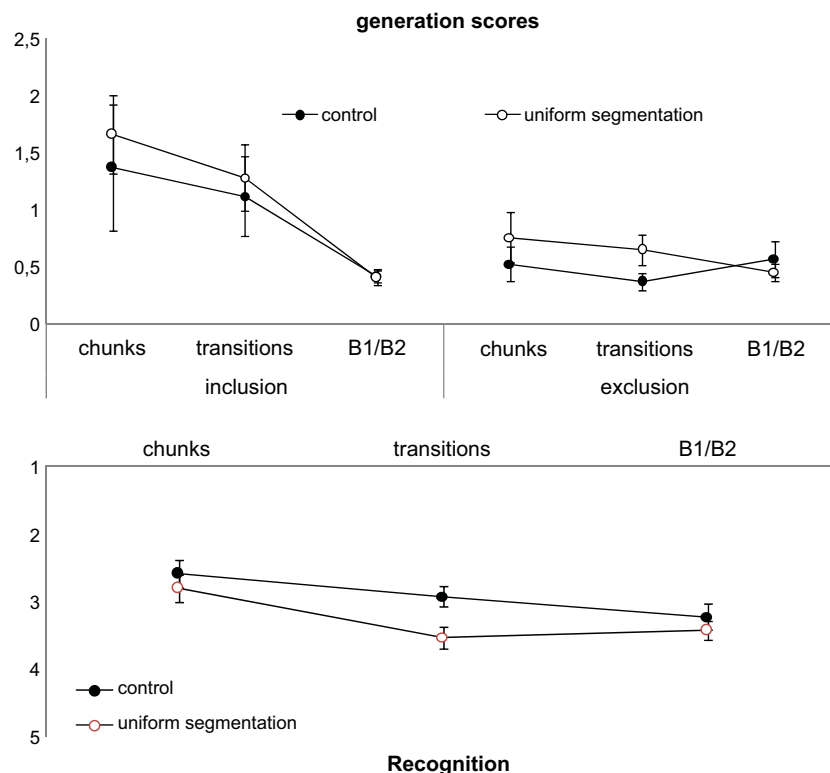


Fig. 5. Direct measures of learning. Upper panel) Average Generation scores for participants in the Control and Uniform Segmentation conditions, represented separately for the inclusion and exclusion tasks, and for those triplets corresponding to chunks, sequential transitions between chunks, or triplets containing a break in an otherwise sequential chunk (B1/B2). Bottom panel) Average Recognition scores for participants in the Control and Uniform Segmentation conditions, and separately for triplets corresponding to chunks, sequential transitions between chunks, or triplets containing a break in an otherwise sequential chunk (B1/B2). Error bars represent Mean Standard Errors.

was not significant either, but there was a significant main effect of Type of triplet, $F(2, 68) = 9.10$, $\eta_p^2 = .21$, $p < .001$. Helmert contrasts showed that chunks were better generated than the other two types of triplets, $F(1, 34) = 8.72$, $\eta_p^2 = .20$, $p < .01$, and that sequential triplets encompassing transitions between chunks were also better generated than non-sequential B1/B2 segments, $F(1, 34) = 9.56$, $\eta_p^2 = .22$, $p < .01$.

As for the recognition task, an ANOVA with Condition and Type of triplet showed a significant effect of Condition, $F(1, 34) = 7.03$, $\eta_p^2 = .17$, $p < .05$, indicating that participants trained with segments associated to systematic colors judged all triplets to be less familiar than did participants in the Control group. There was also a significant effect of Type of triplet, $F(2, 68) = 6.50$, $\eta_p^2 = .16$, $p < .01$, but not a significant interaction Condition \times Type of triplet ($F < 1$). Helmert contrasts indicated that recognition of Chunks was better than that of the other two types, $F(1, 34) = 9.23$, $\eta_p^2 = .21$, $p < .01$, but that there was no difference between other sequential triplets and non-sequential ones ($F < 1$). In sum, the direct measures of sequence learning taken from generation and recognition tasks showed that all participants were, to some extent, able to use sequence knowledge to respond to tasks which explicitly required them to do so. However, responding to specific chunks associated with fixed colors, as arranged in the Uniform Segmentation condition, did not make participants in this group any more able to generate or recognize these chunks than were participants in the Control group. Arguably, the segment knowledge acquired by participants in the Uniform Segmentation group could have been learned associated to the color cues, so that this knowledge was available specifically in the presence of these cues, but was much less accessible when the colors were absent.

3.5. Chunk learning as assessed from the measures of variance

The results presented up to this point suggest that performance in the Uniform Segmentation group was affected by the presence of color cues

associated to the response sequence, but that such segmentation effects were restricted to those blocks in which the colors were present, and they did not survive the removal of the colors over the transfer blocks. An analysis using the proposed ratio between the variance due to responding differently to distinct parts of the sequence, and that due to responding differently over successive repetitions of the same sequence element, reinforced the same conclusion. As can be observed in Fig. 6, the ratio representing the relative impact of between-element variance increased steadily with training in the Uniform Segmentation group, but it remained constant in the Control group. An ANOVA conducted over these scores with Condition as a grouping factor, and with training Blocks (2–5 and 7–8) as a repeated measure showed significant effects of Condition, $F(1, 34) = 26.00$, $\eta_p^2 = .43$, $p < .001$, and Block, $F(5, 179) = 5.13$, $\eta_p^2 = .13$, $p < .001$, as well as a significant interaction between them, $F(5, 179) = 5.79$, $\eta_p^2 = .15$, $p < .001$. In contrast, an analogous ANOVA conducted over the four sequential transfer blocks (9–10 and 12–13) indicated that this pattern of chunked performance did not survive the removal of the color cues (all F s < 1).

A possible account for these results could be taken by arguing that the increase observed in this ratio of variances might not reflect a real fragmentation of sequence knowledge, but rather the fact that in the Uniform Segmentation group the switching between colors was systematically associated to fixed parts of the sequence (i.e., to the first, fourth, seventh and tenth trials). However, if the effect were merely due to the maintenance of a rhythm in the color cues, then a similar ratio should be observed over the test block 6, in which the same pattern of color shifts was maintained, but it was associated with different sequences of responses. As shown in Fig. 6, the effect obtained over the test block 6 for the Uniform segmentation group was not nearly as large as that produced over the training blocks. Consistently, an ANOVA comparing the relevant test block (6) with the average results from the neighboring training blocks (5 and 7) showed a significant effect of type of block, $F(1, 17) = 10.57$, $\eta_p^2 = .38$,

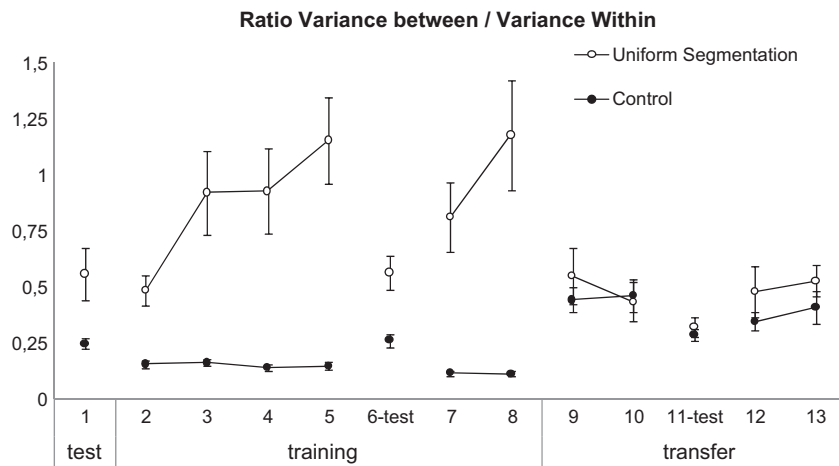


Fig. 6. Chunk learning as assessed from the measures of variance. Average ratio between the variance of performance due to responding differently to each part of the sequence (variance *between elements*) and the variance of performance due to responding differently to repeated presentations of the same element (variance *within elements*) represented separately for participants in the Uniform Segmentation and Control conditions, and for each training and transfer blocks. Error bars represent Mean Standard Errors.

$p < .01$. Plainly, then, systematic color switches were not enough to account for the higher ratios of between-element variance observed over the training blocks. Instead, using colors as cues associated with particular response triplets appears to have facilitated segmented performance, just as it could be expected if participants learned to use the colors as sequence-specific cues. However, training in these conditions was not enough to produce an independent representation of a response chunk which could later sustain discontinuous performance when the colors were removed, and participants were presented with a continuous SRT task.

4. Discussion

The primary aim of this study was to investigate whether learning about a homogeneous sequence in an SRT task proceeds continuously by extracting the statistical regularities arising from the sequence (as proposed for instance in the SRN model, see Cleeremans, 1993) or whether it involves breaking up the sequence into a number of discrete segments, which could later become concatenated into larger chunks, as proposed in Perruchet's PARSE model (Perruchet & Vinter, 1998). These two alternative accounts have been difficult to distinguish from each other, because they both predict the same outcome by the end of training, and because they behave similarly when faced with statistically heterogeneous structures. In contrast, we have argued that the predictions of these two conceptual frameworks diverge when they are applied to learning about a statistically homogeneous structure. In that case, a statistically-based model would predict a rather homogeneous learning effect, spanning simultaneously over the whole sequence, whereas a chunking process would proceed by breaking up the sequence into disjunctive segments, thus resulting in a more discontinuous performance. Yet, given that the segments created by each learner could depend on idiosyncratic factors, fragmentation effects might get obscured when assessed through the typical measures which average performance over learners, and thus result in what may look like a homogeneous learning effect.

In this context, we proposed a measure inspired by the logic of the analysis of variance, and designed an experiment aimed at validating the measure. We proposed that segmentation in learning could be assessed as an increase in the ratio, for each subject and block, between the variance of performance due to responding differently to each part of the sequence (i.e., the variance *between elements*) and the variance of performance attributable to responding differently to repeated presentations of the same element (i.e., variance *within elements*). To take a first step at validating this measure, we trained

two groups of learners with the same SOC sequence, but we induced one of them (the Uniform Segmentation group) to parse the sequence in a uniform way, by using color cues associated to each of the four triplets which make up the whole structure. Conversely, the control group was presented with exactly the same sequence and was prompted by a similar series of color cues, but in this case the series of color cues was decoupled from the response sequence. To assess whether the induced chunks produced an enduring effect on the encoding of the overall sequence, all participants were transferred to a standard SRT task after training, thus requiring them to produce the same sequence of responses without the color cues. We finally assessed learning through direct measures of generation and recognition of segments.

Overall, the results showed that participants in both groups learned the response sequence, and that the manipulation was successful to produce the intended segmentation in performance over the training phase. Participants in the Uniform Segmentation condition responded much faster to the trials internal to a chunk than to those corresponding to the transition between successive chunks. However, the analyses conducted over the transfer phase indicated that the use of response chunks did not survive the removal of the color cues. Indeed, over that transfer phase participants in the Uniform Segmentation group did not respond faster than the control group to those items internal to a chunk, but they tended to respond slower to those items constituting transitions between chunks. Moreover, the proposed score based on the ratio between the relevant components of variance was sensitive enough to capture the group differences in the segmentation patterns observed over the training phase, but it did not reveal any significant difference between groups over the transfer phase.

The pattern of results obtained in this experiment can be taken as supporting the use of the proposed ratio between components of variance as an index of chunking in sequence learning. In contrast, the results are not consistent with the claim that the effects of sequence learning should be taken, in general, as relying on the representation and use of discontinuous segments. Indeed, our results indicate that uniform segments were detected by participants in the Uniform Segmentation condition, that they were recognized to a higher extent than other sequential segments, and that they were used, with the help of the appropriate color cues, to improve RT in response to the relevant sub-sequences. However, the evidence also shows that this knowledge was no longer applied when participants were transferred to a continuous SRT task in which the color cues were no longer available. Thus, it seems as if the knowledge provided by these segments was only useful to the extent that it was supplemented by

salient pointers, such as the global cues, which might have helped participants to retrieve and to hold in memory the relevant information required to trigger the response sequences associated to such cues. In the absence of these color-cues, however, it is remarkable that sequence learning did not completely disappear, but it continued to benefit performance in a more continuous way, resembling the effects observed over the control group. This might indicate that using chunks and statistical learning are independent mechanisms, and that the use of motor chunks relies more on the existence of relatively salient retrieval cues.

As a whole, we surmise that the obtained pattern of results is not consistent with the third scenario described in Perruchet and Pacton (2006), which takes chunk learning as a primary, or even as the single mechanism of sequence learning. In contrast, the results are more consistent with any of the alternative scenarios, which describe chunk learning and statistical learning either as two independent mechanisms, or as two successive steps starting with statistical learning. Our results suggests that chunk learning arises much more clearly when it can be associated with a small series of salient cues, whereas statistical learning can be taken as a background, much more continuous process arising in more implicit settings, and in the absence of such salient cues. We surmise that acknowledging the existence of these two different processes could help us to understand why presenting participants with the same sequence of responses can give place to segmented performance under certain circumstances, but can result in more continuous benefits under other conditions. Thus, in conditions such as those arranged in our Uniform Segmentation group, or in those often used in the DSP tasks (Verwey, 1999, 2001), the use of global cues may help participants to trigger the corresponding series of key-presses in response to such cues with relatively low cognitive cost. In contrast, in the absence of such triggering cues, retrieving the information which would be necessary to disambiguate any given context could be just too demanding for the learner, or may take too much time. In these conditions, therefore, participants would likely resort to a “reaction mode” (Verwey, 2003a, 2003b), thus responding to each trial as it comes, without making any explicit effort to retrieve the fragmentary knowledge that they may have associated to specific cues. Performance in this case would not reflect such fragmentary knowledge, but only the more continuous, statistical, and implicit knowledge accrued simultaneously from practice with that task.

The conclusions that (1) chunk knowledge is separable from the continuous, statistically-based, and implicit knowledge which arises from performing a sequence learning task, and that (2) the use of chunk knowledge requires the aid of some salient, disambiguating cues, are both consistent with previous results such as those reported in Verwey (2003a, 2003b, 2010) and Kirsch et al. (2010). Verwey (2003a, 2003b) showed that participants trained with a DSP task containing sequence-specific cues failed to apply the acquired knowledge to respond to a subsequent SRT task which included the previously trained sequences, now presented continuously and without the training cues. In addition, the pattern of response times found with elderly in the DSP task suggests that elderly do not use motor chunks during practice, but still show improvements which can be explained by statistical learning (Verwey, 2010, see also Verwey et al., in press). Our results confirmed, and go beyond these observations, in showing that a similar lack of transfer of chunk learning can be obtained by removing the global cues, even without masking the persistence of the trained sequence. Thus, this evidence can be regarded as supporting Verwey's (2003a, 2003b) framework on the existence of independent processors which race against each other to trigger each next movement in a familiar movement sequence. Moreover, we contend that the same conclusion is consistent with some recent results reported by Kirsch et al. (2010) (see also Koch & Hoffmann, 2000), indicating that certain salient patterns can contribute to the expression of chunk learning, even in a continuous SRT task. In addition to producing some effects independent of learning

(Jiménez, 2008), the use of salient, ascending and descending runs may well facilitate retrieval of the relevant knowledge, and the use of this knowledge in triggering the appropriate series of responses, thus decreasing the cost associated with the explicit use of such fragmentary knowledge. However, even if this is possible under such specific conditions, their exceptional nature indicates that the role of such fragmentary knowledge might remain relatively minor in accounting for the learning observed in a continuous SRT task, whereas a larger part of these effects could rely on more continuous processes of associative, or statistically-based, accrual of sequence information.

Acknowledgments

This research was supported by the Spanish *Ministerio de Educación y Ciencia* (MEC) with research grants SEJ2005-25754-E, and SEJ2006-27564 to Luis Jiménez. We wish to thank Axel Cleeremans for his continuous insight, and specifically for his suggestion on using the opposite instructions to make the exclusion generation task more comparable with the inclusion task. We wish to thank Pierre Perruchet and Dennis Rüniger for their helpful suggestions on a previous version of the manuscript.

Appendix A

A.1. Definition

Let's consider a matrix A compounded of $m \times n$ elements of type real. Its mean \bar{X} can be calculated as following:

$$\bar{X} = \frac{\sum_{j=1}^m \sum_{i=1}^n x_{j,i}}{m \cdot n} = \frac{\sum_{j=1}^m \bar{X}_j}{m}$$

with $x_{j,i}$ element at j th column and i th line in A; \bar{X}_j mean of the j th column.

The estimation of variance of A is defined by:

$$var(A) = \frac{\sum_{j=1}^m \sum_{i=1}^n (x_{j,i} - \bar{X})^2}{m \cdot n - 1}$$

Whereas we introduce two new measures of variance, respectively:

$$var_{Within}(A) = \frac{\sum_{j=1}^m \sum_{i=1}^n (x_{j,i} - \bar{X}_j)^2}{m \cdot (n - 1)} \text{ and } var_{Between}(A) = \frac{\sum_{j=1}^m (\bar{X}_j - \bar{X})^2}{m - 1}$$

with A matrix of dimensions $m, n \in N^*$, $x_{j,i}$ element at j th column and i th line in A, \bar{X}_j mean of the j th column of A, and \bar{X} global mean of A.

A.2. Equations

Our aim in the following demonstration is to find an equation linking $var(A)$, $var_{Within}(A)$ and $var_{Between}(A)$ together. Thus, starting from:

$$var(A) = \frac{\sum_{j=1}^m \sum_{i=1}^n (x_{j,i} - \bar{X})^2}{m \cdot n - 1}$$

We extrapolate the sum:

$$\begin{aligned} (m \cdot n - 1) \cdot var(A) &= \sum_{j=1}^m \sum_{i=1}^n [(x_{j,i} - \bar{X}_j) + (\bar{X}_j - \bar{X})]^2 \\ &= \sum_{j=1}^m \sum_{i=1}^n [(x_{j,i} - \bar{X}_j)^2 + (\bar{X}_j - \bar{X})^2 + 2 \cdot (x_{j,i} - \bar{X}_j)(\bar{X}_j - \bar{X})] \\ &= \sum_{j=1}^m \sum_{i=1}^n (x_{j,i} - \bar{X}_j)^2 + \sum_{j=1}^m \sum_{i=1}^n (\bar{X}_j - \bar{X})^2 + 2 \cdot \sum_{j=1}^m \sum_{i=1}^n (x_{j,i} - \bar{X}_j)(\bar{X}_j - \bar{X}). \end{aligned}$$

Then, considering the last term alone:

$$\begin{aligned} \sum_{j=1}^m \sum_{i=1}^n (x_{j,i} - \bar{X}_j) (\bar{X}_j - \bar{X}) &= \sum_{j=1}^m \sum_{i=1}^n (x_{j,i} \cdot \bar{X}_j - \bar{X}_j^2 - x_{j,i} \cdot \bar{X} + \bar{X}_j \cdot \bar{X}) \\ &= \sum_{j=1}^m \sum_{i=1}^n (x_{j,i} \cdot \bar{X}_j - \bar{X}_j^2) - \sum_{j=1}^m \sum_{i=1}^n x_{j,i} \cdot \bar{X} \\ &\quad + \sum_{j=1}^m \sum_{i=1}^n \bar{X}_j \cdot \bar{X} \\ &= \sum_{j=1}^m \left[\sum_{i=1}^n x_{j,i} \cdot \bar{X}_j - \sum_{i=1}^n \bar{X}_j^2 \right] - \bar{X} \cdot \sum_{j=1}^m \sum_{i=1}^n x_{j,i} \\ &\quad + \bar{X} \cdot \sum_{j=1}^m \sum_{i=1}^n \bar{X}_j. \end{aligned}$$

Since \bar{X}_j is constant over i , we obtain:

$$\begin{aligned} &= \sum_{j=1}^m \left[\bar{X}_j \cdot \sum_{i=1}^n x_{j,i} - n \cdot \bar{X}_j^2 \right] - \bar{X} \cdot (m \cdot n \cdot \bar{X}) + \bar{X} \cdot (m \cdot n \cdot \bar{X}) \\ &= \sum_{j=1}^m \left[\bar{X}_j \cdot (n \cdot \bar{X}_j) - n \cdot \bar{X}_j^2 \right] = \sum_{j=1}^m \left[n \cdot \bar{X}_j^2 - n \cdot \bar{X}_j^2 \right] = 0. \end{aligned}$$

We move back to our equation:

$$\begin{aligned} (m \cdot n - 1) \cdot \text{var}(A) &= \sum_{j=1}^m \sum_{i=1}^n (x_{j,i} - \bar{X}_j)^2 + \sum_{j=1}^m \sum_{i=1}^n (\bar{X}_j - \bar{X})^2 + 0 \\ \text{var}(A) &= \frac{\sum_{j=1}^m \sum_{i=1}^n (x_{j,i} - \bar{X}_j)^2}{m \cdot n - 1} + \frac{\sum_{j=1}^m \sum_{i=1}^n (\bar{X}_j - \bar{X})^2}{m \cdot n - 1}. \end{aligned}$$

We may also write:

$$\text{var}(A) = \frac{m \cdot (n-1)}{m \cdot (n-1)} \cdot \frac{\sum_{j=1}^m \sum_{i=1}^n (x_{j,i} - \bar{X}_j)^2}{m \cdot n - 1} + \frac{m-1}{m-1} \cdot \frac{n \cdot \sum_{j=1}^m (\bar{X}_j - \bar{X})^2}{m \cdot n - 1}.$$

And finally we obtain:

$$\text{var}(A) = \frac{m \cdot (n-1)}{m \cdot n - 1} \cdot \text{var}_{\text{Within}}(A) + \frac{n \cdot (m-1)}{m \cdot n - 1} \cdot \text{var}_{\text{Between}}(A).$$

References

Bo, J., & Seidler, R. D. (2009). Visuospatial working memory capacity predicts the organization of acquired explicit motor sequences. *Journal of Neurophysiology*, 101, 3116–3125.

Christiansen, M. H., Allen, J., & Seidenberg, M. S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes*, 13, 221–268.

Cleeremans, A. (1993). *Mechanisms of implicit learning: Connectionist models of sequence processing*. Cambridge, MA: MIT Press.

Cleeremans, A., & Jiménez, L. (1998). Implicit sequence learning: The truth is in the details. In M. Stadler, & P. Frensch (Eds.), *Handbook of implicit learning* (pp. 323–364). Thousand Oaks, CA: Sage.

Cock, J., Berry, D., & Buchner, A. (2002). Negative priming and sequence learning. *European Journal of Cognitive Psychology*, 14, 27–48.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review*, 8, 343–350.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179–211.

Jiménez, L. (2008). Taking patterns for chunks: Is there any evidence of chunk learning in continuous serial reaction-time tasks? *Psychological Research*, 72, 387–396.

Jiménez, L., & Méndez, C. (1999). Which attention is needed for implicit sequence learning? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 236–259.

Jiménez, L., Méndez, C., & Cleeremans, A. (1996). Comparing direct and indirect measures of sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 948–969.

Kennerley, S. W., Sakai, K., & Rushworth, M. F. S. (2004). Organization of action sequences and the role of pre-SMA. *Journal of Neurophysiology*, 91, 978–993.

Kirsch, W., Sebald, A., & Hoffmann, J. (2010). RT patterns and chunks in SRT tasks: A reply to Jiménez (2008). *Psychological Research*, 74, 352–358.

Koch, L., & Hoffmann, J. (2000). Patterns, chunks, and hierarchies in serial reaction-time tasks. *Psychological Research*, 63, 22–35.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirement of learning: Evidence from performance measures. *Cognitive Psychology*, 19, 1–32.

Perruchet, P., & Tzelgov, J. (2009). Automatic mode of acquisition of spatial sequences in a serial reaction time paradigm. *Psychological Research*, 73, 98–106.

Perruchet, P., & Gallego, J. (1997). A subjective unit formation account of implicit learning. In D. Berry (Ed.), *How implicit is implicit learning?* (pp. 124–161). Oxford: Oxford University Press.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, 10, 233–238.

Perruchet, P., & Vinter, A. (1998). Learning and development: The implicit knowledge assumption reconsidered. In M. A. Stadler, & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 495–531): Sage Publications.

Perruchet, P., & Vinter, A. (2002). The self-organized consciousness. *The Behavioral and Brain Sciences*, 25, 297–388.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6, 855–863.

Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 585–594.

Rünger, D., & Frensch, P. A. (2008). How incidental sequence learning creates reportable knowledge: The role of unexpected events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 1011–1026.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, 35, 606–621.

Stadler, M. A. (1993). Implicit learning: Questions inspired by Hebb (1961). *Memory & Cognition*, 21, 819–827.

Stadler, M. A. (1995). Role of attention in implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 674–685.

Verwey, W. B. (1999). Evidence for a multi-stage model of practice in a sequential movement task. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 1693–1708.

Verwey, W. B. (2001). Concatenating familiar movement sequence. The versatile cognitive processor. *Acta Psychologica*, 106, 69–95.

Verwey, W. B. (2003a). Processing modes and parallel processors in producing familiar keying sequences. *Psychological Research*, 67, 106–122.

Verwey, W. B. (2003b). Effect of sequence length on executing familiar keying sequences: Lasting segmentation and preparation? *Journal of Motor Behavior*, 35, 343–354.

Verwey, W. B. (2010). Diminished motor skill development in elderly: Indications for limited motor chunk use. *Acta Psychologica*, 134, 206–214.

Verwey, W. B., Abrahamse, E. L., & Jiménez, L. (2009). Segmentation of relatively short keying sequences does not transfer to other sequences. *Human Movement Science*, 28, 348–361.

Verwey, W. B., Abrahamse, E. L., Ruitenberg, M.F.L., Jiménez, L. & de Kleine, E. (in press). Motor skill learning in the middle-aged: limited development of motor chunks and explicit sequence knowledge, *Psychological Research*, doi:10.1007/s00426-011-0320-0.

Verwey, W. B., & Eikelboom, T. (2003). Evidence for lasting sequence segmentation in the discrete sequence-production task. *Journal of Motor Behavior*, 35, 171–181.

Wilkinson, L., & Shanks, D. R. (2004). Intentional control and implicit sequence learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 354–369.

Ziessler, M., & Nattkemper, D. (2001). Learning of event sequences is based on response-effect learning: Further evidence from a serial reaction task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 595–613.