Grouping in working memory guides chunk formation in long-term memory: Evidence from the Hebb effect

Musfeld, Philipp ; Dutli, Joscha ; Oberauer, Klaus ; Bartsch, Lea M

DOI: https://doi.org/10.1016/j.cognition.2024.105795

Originally published at:
Musfeld, Philipp; Dutli, Joscha; Oberauer, Klaus; Bartsch, Lea M (2024). Grouping in working memory guides chunk formation in long-term memory: Evidence from the Hebb effect. Cognition, 248:105795.
DOI: https://doi.org/10.1016/j.cognition.2024.105795

Cognition 248 (2024) 105795

Contents lists available at ScienceDirect

Cognition

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

Grouping in working memory guides chunk formation in long-term memory: Evidence from the Hebb effect

Philipp Musfeld *, Joscha Dutli, Klaus Oberauer, Lea M. Bartsch

University of Zurich, Department of Psychology, Cognitive Psychology Unit, Binzmühlestrasse 14, Box 22, CH-8050 Zurich

ARTICLE INFO

Keywords:
Working memory
Long-term memory
Repetition learning
Chunking
Hebb effect

ABSTRACT

The Hebb effect refers to the improvement in immediate memory performance on a repeated list compared to unrepeated lists. That is, participants create a long-term memory representation over repetitions, on which they can draw in working memory tests. These long-term memory representations are likely formed by chunk acquisition: The whole list becomes integrated into a single unified representation. Previous research suggests that the formation of such chunks is rather inflexible and only occurs when at least the beginning of the list repeats across trials. However, recent work has shown that repetition learning strongly depends on participants recognizing the repeated information. Hence, successful chunk formation may depend on the recognizability of the repeated part of a list, and not on its position in the list. Across six experiments, we compared these two alternatives. We tested immediate serial recall of eight-letter lists, some of which partially repeated across trials. We used different partial-repetition structures, such as repeating only the first half of a list, or only every second item. We manipulated the salience of the repeating structure by spatially grouping and coloring the lists according to the repetition structure. We found that chunk formation is more flexible than previously assumed: Participants learned contiguous repeated sequences regardless of their position within the list, as long as they were able to recognize the repeated structure. Even when the repeated sequence occurred at varying positions over repetitions, learning was preserved when the repeated sequence was made salient by the spatial grouping. These findings suggest that chunk formation requires recognition of which items constitute a repeating group, and demonstrate a close link between grouping of information in working memory, and chunk formation in long-term memory.

1. Introduction

Repetition learning plays a fundamental role in acquiring new knowledge and skills. For example, it is a common practice to engage in repeated study or exposure to content until we have solidly memorized something for an exam. This iterative process of repetition learning involves both working memory and long-term memory (Burgess & Hitch, 2005; Page & Norris, 2009). Working memory is a capacity limited system and is needed to hold currently studied information temporally available for use in thought and action (Cowan, 2017; Luck & Vogel, 1997; Oberauer, 2009). Long-term memory serves as a repository for our knowledge and experiences and has an extensive storage capacity (Brady et al., 2008; Tulving, 1972).

A well-known example to study repetition learning experimentally is the Hebb paradigm. In the Hebb paradigm, participants are presented with several memory lists for an immediate memory test. One of these lists, the *Hebb list*, is repeated occasionally. What is typically observed is that memory performance improves with repetitions for the repeated list but not for the non-repeated lists (Hebb, 1961). This finding demonstrates that repeated exposure to the same information in working memory can lead to the formation of stable representations in long-term memory, which, in turn, can be used to assist performance in a working memory task (Burgess & Hitch, 2005; Mızrak & Oberauer, 2022; Page & Norris, 2009; Souza & Oberauer, 2022). Although the Hebb effect has been demonstrated for a wide range of different materials (Couture & Tremblay, 2006; Johnson & Miles, 2019; Musfeld et al., 2023b; Souza & Oberauer, 2022; Sukegawa et al., 2019), it has been most extensively used to study verbal sequence learning (i.e., sequences of digits, letters, or phonemes). The Hebb effect for verbal materials has even been proposed as a model for the acquisition of new word forms (Norris et al.,

* Corresponding author.
E-mail addresses: philipp.musfeld@psychologie.uzh.ch (P. Musfeld), joscha.dutli@psychologie.uzh.ch (J. Dutli), k.oberauer@psychologie.uzh.ch (K. Oberauer), l.bartsch@psychologie.uzh.ch (L.M. Bartsch).

https://doi.org/10.1016/j.cognition.2024.105795
Received 31 January 2024; Received in revised form 5 April 2024; Accepted 15 April 2024
Available online 25 April 2024

2018; Page & Norris, 2009; Saint-Aubin & Guérard, 2018; Szmalec et al., 2009). Yet, it is not clear how the information learned through repetition is represented in long-term memory, and how these representations can benefit working memory. Here, we aim to address this question.

1.1. Theories on the mechanisms underlying Hebb repetition learning

Originally, researchers have approached this question by considering how serial order of information within a sequence is coded in memory. Two major families of models have been proposed: The first are *chaining models*, which are based on the idea that a sequence is coded by forming associations between subsequent items. When retrieving a sequence, each item serves as a cue for retrieving the next item, thereby reconstructing the sequence in its correct order (Lewandowsky & Murdock, 1989; Murdock, 1995; Solway et al., 2012). Although these models have been largely abandoned due to their inability to account for several empirical benchmark findings in serial recall (Henson, 1999; Henson et al., 1996; Oberauer et al., 2018; Osth & Dennis, 2015; Osth & Hurlstone, 2023), variations of the chaining idea still build the foundation of successful computational models of serial order (Logan, 2021; Logan & Cox, 2021). The second family of models are *positional models,* which are based on the idea that a sequence is coded by associating each item to its relative position within the list. At test, the positional context of an item serves as a cue for retrieving the item (Brown et al., 2000; Cumming et al., 2003; Farrell, 2006; Henson, 1999). Positional models are far more widely accepted than chaining models and build the foundation of several successful models of short-term and working memory (Brown et al., 2007; Farrell, 2012; Lewandowsky & Farrell, 2008; Oberauer et al., 2012).

Within the research on the Hebb effect, both models of serial order provide grounds for explaining how the representation of a repeated sequence becomes stronger over repetitions, but for different reasons: In a chaining account of repetition learning, it is plausible that repetitions strengthen the association between subsequent items, whereas in a positional account of repetition learning, repetitions plausibly strengthen the association between an item and its position within the list.

Hitch et al. (2005) tested the explanatory power of these accounts of repetition learning by making use of a Hebb-like paradigm in which only specific parts of a sequence repeated over trials. In their experiments, participants encoded lists of letters of various lengths for an immediate serial recall test. Within blocks of four consecutive trials, specific parts of the lists were repeated. To test the predictions of the chaining account of learning, the authors constructed lists which always contained the same partial sequence of letters (e.g. “BZL”) but at different positions within the list (i.e., shifting a coherent partial sequence within the whole list, “ST**BZL** – **BZL**GH - **ZL**JH**B**”). This condition solely repeated associations between subsequent items, but not between the items and their position within the list. To compare this to the predictions of the positional account of repetition learning, the authors constructed lists in which only every other item repeated (“**B**S**Z**F**L** – **B**M**Z**N**L** – **B**C**Z**H**L**”). This condition solely repeated associations between an item and its position within the list, without repeating any associations between subsequent items. Additionally, the authors also realized conditions in which either the beginning or the end of a sequence repeated. These conditions repeated both item-item and item-positions associations, but in one case positions are repeated relative to the beginning of the list and in the other, relative to the end of the list. Their results showed learning effects through repetition only when the beginning of the sequence repeated, but in none of the other conditions. These results are inconsistent with predictions from both a chaining and a positional account of repetition learning, suggesting that learning of sequences can be explained by neither of the two (see also Cumming et al. (2003), and Schwartz and Bryden (1971), for similar conclusions).

Hitch et al. (2005) proposed that learning of repeated sequences might involve more global processes like chunk formation, which operates on the level of a sequence as a whole, instead of the level of single item associations. Chunk formation is understood as a process by which multiple separate elements of information are integrated into one unified representations of these elements and their structure (Cowan & Chen, 2008; Ericsson & Kintsch, 1995; Gobet et al., 2016; Miller, 1956), thereby allowing to represent the same amount of information in a more efficient (or compressed) way (Brady et al., 2009; Chekaf et al., 2016; Huang & Awh, 2018; Norris et al., 2020; Norris & Kalm, 2021; Thalmann et al., 2019). In sequence learning, this could be the integration of a string of letters into an acronym, or the integration of syllables into a word. Chunk formation in itself is a rather broad concept, and the mechanisms underlying the associated efficiency gain in representing information are still debated (see, e.g., Gobet et al. (2016) or Norris and Kalm (2021) for reviews). Yet, it has been widely accepted as a fundamental mechanism for learning of new representations (Anderson, 1993; Burgess & Hitch, 2005, 2006; French et al., 2011; Gobet & Simon, 1996; Huang & Awh, 2018; Jones et al., 2014; Mızrak & Oberauer, 2022; Orbán et al., 2008; Page & Norris, 2009; Robinet et al., 2011).

1.2. The role of awareness in repetition learning and chunk formation

Over the past decades, repetition learning has often been considered an example for an implicit learning process, that is, people don’t have to be aware that they are encoding the same information over and over again for building up a long-term memory representation (Attout et al., 2020; Couture & Tremblay, 2006; Guérard et al., 2011; Hebb, 1961; McKelvie, 1987). Yet, mechanistic explanations of repetition learning have agreed that some form of recognition mechanism is required to ensure that long-term memory representation of previous experiences are used only, when a sufficient overlap between a previous and a new encounter of the same information is detected (Burgess & Hitch, 2005, 2006; Page & Norris, 2009). More recently, Musfeld et al. (2023a) provided strong empirical support that recognizing a repeated memory list is indeed a necessary condition for repetition learning, and that such a recognition process does not occur implicitly. That study tracked learning on the level of individual participants and showed that participants only learn after they explicitly recognized the repetition of a repeating memory list. Until the repetition was recognized, no learning was observed (see also Ngiam et al., 2019, for similar findings in associative learning tasks).

The insight that recognition of repeating patterns is a necessary condition for learning also offers an explanation for why partial repetitions lead to learning at the beginning of a list but not at other positions in the list: It makes them easier to be recognized. One can think about it as follows: In a standard Hebb paradigm, in which a whole list is repeated, the information which constitute a repeating unit (i.e. the entire repeated list) is clearly defined by the start and the end of the list. However, with partially repeated list structures like the ones used by Hitch et al. (2005), it is harder to recognize which subset of the items could be learned as a new unit. It might still be relatively easy to identify the repeated information if it is presented at the beginning of the list, as it matches the natural boundary of the list. However, moving the repeated information to the end of the list, or even interleaving it with other items, might make it more difficult to recognize the repetition within the sequence, thereby preventing learning in the first place.

To account for these differences in the ability to recognize and learn repeated patterns in partially repeated lists, Burgess and Hitch (2006) proposed a cumulative recognition mechanism, which operates incrementally from the beginning of a list. In their model, every presented list becomes associated with a context set that is stored in long-term memory. When a new list is presented, the overlap between the encoded list and all available context sets in long-term memory is computed. The overlap is computed incrementally as the list is presented: After every new item the cumulative mismatch of the new list up to that item is computed for every context set in long-term memory. Every context set for which the degree of mismatch exceeds a threshold is discarded. Thus, when partially repeated lists don’t match at the beginning, all context sets will be discarded and the repetition in a later section of the list will never be recognized, preventing learning.

The cumulative recognition mechanism proposed by Burgess and Hitch (2006) helps to account for the findings observed by Hitch et al. (2005), but also introduces rather strong constraints on the flexibility with which repeated patterns can be recognized (and learned) more generally. Following their model, repetitions are only recognizable and beneficial for learning when they occur at the beginning of a sequence. However, the recognizability of a repeated segment arguably not only depends on its position within a sequence but might also be influenced by other variables. Hence, people might be far more flexible in learning repeated patterns, as long as they are able to recognize them. We argue that if we increase the recognizability of a repeated pattern within a list by making it more salient, people should be able to learn these patterns regardless from their position within a sequence. However, if chunk formation in sequence learning is indeed limited by positional constraints, facilitating participants' awareness of repeated list structures should not help learning. The goal of the present study is to test this hypothesis.

2. The present study

The present study provides a conceptual replication of the study by Hitch et al. (2005), while also extending it by an experimental manipulation of the recognizability of the repeated structure within the lists. We conducted six experiments, which realized six different partially repeated list conditions. We used a typical Hebb paradigm, in which participants saw sequences of eight consonants for an immediate serial recall task. On every third trial, half of the presented consonants were repeated according to the partial repetition structure of the experiment (these lists are referred to as the *partial Hebb lists*). We used the following repetition structures: *start repeat*, repeating the items at the first four list positions (Experiment 1); *middle repeat*, repeating the items at the four positions in the middle of the list (Experiment 2); *end repeat*, repeating the items at the last four positions of the list (Experiment 3); *chaining,* repeating a coherent string of four subsequent consonants, but at varying positions within the list (Experiment 4); *odd repeat,* repeating the items at the odd positions of the list (Experiment 5); *even repeat,* repeating the items at the even positions of the list (Experiment 6).

For all experiments, we manipulated the recognizability of the repeated structure within the list in a between-subjects design. In the *No Salience* condition, participants encoded all eight consonants sequentially across a row of eight black boxes. This served as a control condition, in which the potentially repeated list structure was inconspicuously integrated into the rest of the list; we expected this condition to replicate the findings by Hitch et al. (2005). In the *High Salience* condition, we increased the salience (and therefore the recognizability) of the repeated structure by spatially grouping and coloring the eight boxes in accordance with the repetition structure of the experiment (see Fig. 1 for a visualization of all different highlighting schemes). Crucially, the highlighting scheme was not only applied to the partial Hebb list but also to all unrepeated Filler lists, to not make the Hebb list stand out against the Filler lists, but to only highlight the potentially repeated part within the presented lists. Our prediction was that the highlighting of the repeated structure would facilitate participants' awareness of which items can be integrated into a chunk, and as a result, enable learning even in those conditions that did not lead to learning in the study by Hitch et al. (2005).

In the following, we present the results of the six experiments in three parts. In the first part, we report the results of the *start repeat* (Experiment 1), *middle repeat* (Experiment 2) and *end repeat* (Experiment 3) conditions, asking whether the repeated items have to be presented at the beginning of the list for learning to occur. These experiments clearly establish that learning of the repeated sequence was not limited to conditions in which the repetition occurred at the beginning of the list. Partial repetitions could also be learned in the middle or at the end of the list. Yet, these experiments only showed weak effects of our salience manipulation. This could mean that, contrary to our hypothesis, the recognizability of the repeated sequence did not affect learning, or else, that the sequences were relatively easy to recognize even without additional highlighting. We address this issue in the second part with the results of the *chaining* condition (Experiment 4). This experiment revealed strong effects of the salience manipulation, indicating that when it was hard to recognize the repeated sequence within the list, learning could be facilitated by increasing its recognizability. In the last part, we present the results from the *odd repeat* (Experiment 5) and *even repeat* (Experiment 6) conditions to further test the flexibility of chunk formation under conditions in which the repeated information is interleaved by other items.

Overall, our results reveal that people can learn repeated sequences much more flexibly than suggested in previous work. Learning was not dependent on any specific list condition, but on (1) participants' ability to recognize the repeated structure within the list, and (2) the repeated items being presented as a coherent string. We propose that these results are best explained by assuming a direct link between the grouping of information in working memory, and the formation of new chunks in long-term memory: How information is structured in working memory during encoding guides chunk formation in long-term memory.

3. General method

3.1. Transparency and openness

All experiments in this study were preregistered on the Open Science Framework (OSF) prior to data collection, and all data analyses were conducted as described in our preregistered analysis plan. The preregistrations for Experiments 1–3 and 5–6 are available at https://osf.io/dgu2z and the preregistration for Experiment 4 is available at https://osf.io/8dg5v.[1] Experimental software, data, and analysis scripts for all experiments reported here are available in the OSF at https://osf.io/pb6vx/ (Musfeld, Dutli, Oberauer, & Bartsch, 2023) https://osf.io/pb6vx/?view_only=e4bcd9fb131e49eb97544a654d6e8d7d.

The experiments were programmed using the free and open online experiment builder *lab.js* (Henninger et al., 2022). Data analyses were conducted using *R* (R Core Team, 2023) and the R-packages *tidyverse* (Wickham et al., 2019), *here* (Müller, 2020), *rstan* (Stan Development Team, 2023)*, brms* (Bürkner, 2017), *tidybayes* (Kay, 2023), and *bayestestR* (Makowski et al., 2019).

3.2. Participants and exclusion criteria

All participants were recruited online, either from the student population of the University of Zurich or from the online participant platform *Prolific*. For Experiments 1–3 and 5–6, data collection was part of a university course and an initial sample from the student population of the University of Zurich participated in exchange for partial course credit. We recruited additional participants on *Prolific*, as evidence for our research questions remained inconclusive in the initial sample. For Experiment 4, we recruited all data on *Prolific*. Participants on *Prolific* received £4.50 for their participation. Final sample sizes and demographic information for all Experiments are presented in Table 1.

All participants were between 18 and 35 years old, fluent in German, and did not have any language or speech related disorder. They were not allowed to participate in more than one of the Experiments. To ensure high data quality, several quality checks were implemented: First, participants had to complete a questionnaire after reading the instructions, which tested their understanding of the study. Participants were only

[1] We originally conducted the Experiments in a different order but changed it for the paper to create a more coherent structure. There is no deviation from the preregistered hypotheses we tested with the different experiments.

Fig. 1. Illustration of the general experimental design and the highlighting of specific list structures in the High Salience Condition of the six experiments. The panel in the first row shows an outline of the general repetition scheme within all six experiments. The two panels in the second row illustrate item presentation in the No Salience Condition and the recall procedure, which was the same for all six experiments. The remaining panels show the highlighting schemes in the High Salience condition for the different list structures implemented in the six experiments.

Table 1
Sample Size and Demographic Information for all Experiments in this Study.

	Before Exclusion		After Exclusion				
Experiment	*N* Collected	*n* Excluded	*N* Final	*n* Prolific	*n* No Salience	*n* High Salience	Age
Start Repeat	122	1	121	43	60	61	18–35 (M = 23.4; SD = 3.67)
Middle Repeat	120	2	118	49	60	58	18–35 (M = 24.9; SD = 4.66)
End Repeat	122	4	118	43	59	59	18–34 (M = 23.0; SD = 3.65)
Chaining	121	1	120	120	60	60	18–35 (M = 25.4; SD = 4.40)
Odd Repeat	118	0	118	40	59	59	18–35 (M = 23.0; SD = 4.14)
Even Repeat	121	3	118	52	58	60	18–35 (M = 23.7; SD = 4.26)

allowed to participate if they answered all questions correctly. Second, we tracked participants' active browser window during the experiment and aborted the study if participants left the browser window during the experiment more than five times. Third, we tracked a measure of careless responding, which was defined as responding with the same letter more than four times within a trial (e.g., K K K K K L L L; it was not possible for a letter to appear twice within a list) and aborted the study if careless responding was detected more than 5 times. Fourth, after completing the experiment, we asked participants if they participated seriously and if they had used any aids to improve their performance during the experiment (Aust et al., 2013). Participants were only considered for data analysis if they indicated serious participation and no use of any aids. Lastly, we excluded all participants from the data analysis whose average performance in Filler trials was at chance level during at least one half of the experiment. Chance level was defined as performance below the 99% quantile of the binomial distribution, with guessing probability set to 1/18 and number of responses set to 160, leading to a cut-off value of at least 10% correct responses.

The experiments were carried out in accordance with the guidelines of the Ethics Committee of the Faculty of Arts and Social Sciences at the University of Zurich. As the experiments involved minimal risk, no formal approval was required. All participants took part after giving informed consent.

3.3. Design

All experiments followed the structure of a typical Hebb paradigm, which is schematically presented in Fig. 1. On each trial, participants were sequentially presented with lists of 8 consonants for an immediate serial recall test and asked to type in the presented consonants in forward order immediately after presentation. In one of these lists, the "partial Hebb list", half of the presented consonants was repeated according to the repetition structure of the corresponding experiment. This partial Hebb list was presented on every third trial, starting with trial two. All other trials consisted of Filler lists, which did not contain any repeating item structures. In total, participants completed 60 trials, leading to 20 repetitions of the partial Hebb list.

Over the six Experiments, we realized six different partially repeated list structures (see Fig. 1 for an overview): (1) In the *Start Repeat* Experiment, the first four consonants of the list repeated; (2) in the *Middle Repeat* Experiment, the four consonants in the middle of the list repeated; (3) in the *End Repeat* Experiment, the last four consonants of the list repeated; (4) in the *Chaining* Experiment, a sequence of four consecutive items repeated over the Experiment, but the position of this sequence was shifted within the list across repetitions. Specifically, the repeated sequence could start somewhere between list positions one and five. Each starting position of the repeated sequence was realized equally often but the order of starting positions throughout the experiment was randomized; (5) in the *Odd Repeat* Experiment, the four items at the odd numbered list positions repeated; (6) in the *Even Repeat* Experiment, the four items at the even numbered list positions repeated.

In all experiments, participants were randomly assigned to one of two between-subject Salience conditions upon starting the experiment: In the *No Salience Condition*, all consonants were presented sequentially in a row of eight black framed boxes in the middle of the screen. There was no additional highlighting of the repeating list structure to increase its salience. In the *High Salience* condition, the consonants were also presented in a row of 8 boxes, but the boxes were spatially grouped in accordance with the repeated list structure of the experiment and the boxes of the repeating list positions were framed in orange. This highlighting was applied to both Filler and Hebb lists to ensure that it only increased the salience of the repeating structure, but not of the partial Hebb list itself. An overview of the highlighting schemes for the different list structures is presented in Fig. 1.

3.4. Stimuli

All memory lists were created randomly for each participant upon the start of the experiment by sampling 8 consonants from the pool of all consonants, excluding "Y" and "W". The following constraints were imposed on the creation of lists: (1) No letter was allowed to appear twice within the same list; (2) the letters "M" and "N" were not allowed to be part of the same list due to their high amount of phonological similarity; (3) all partial Hebb lists had to differ in at least 3 items to avoid item repetition within the unrepeated part of the Hebb lists; (4) all filler lists had to differ in at least 5 item-position associations from all other lists to decrease chances of accidental repetitions; (5) we created a list of 56 well-known German 3- and 2-letter acronyms. These acronyms were not allowed to be part of a list to avoid any unintended chunking effects (the list of known acronyms is available in the OSF).

3.5. Procedure

All participants took part online from their own devices (computers only, no tablets or phones). The experiment began with a detailed explanation of the task. Participants were not informed about the possibility of repetitions within lists. After reading the instructions, participants had to complete a short test on their understanding of the task, which they had to pass in order to take part in the experiment.

After passing the instruction test, participants performed three practice trials to make themselves familiar with the experimental task, before moving on to the main part of the study. Each trial of the experiment started with the presentation of 8 empty boxes for 1000 ms, which were spatially aligned and colored according to the repetition scheme of the experiment and the assigned salience condition (see Fig. 1). Afterwards, the 8 consonants appeared sequentially from left to right inside their corresponding boxes. Each consonant was visible for 500 ms, with a 100 ms inter-stimulus-interval.[2] After the last consonant had been presented, an immediate serial recall task followed without an additional retention interval. For this, the 8 boxes stayed on screen and a prompter indicated to participants to type in the letters into their corresponding box in forward order. In the *No Salience* condition, the recall phase looked the same as the presentation phase. In the *High Salience* condition, the spatial alignment and coloring of the boxes was removed so that the recall phase looked the same as for the *No Salience* condition (see Fig. 1). After recalling all consonants, participants received short feedback about the number of correctly recalled items (e.g. "5/8 answers were correct!") and moved on to the next trial at a self-chosen pace.

After finishing all 60 trials of the serial recall task, a short questionnaire followed in which participants were asked about their awareness of the repeated list structure: First, participants were asked if they had noticed anything special about the experimental design and typed in their response in an open-ended text field. Next, participants were informed that there was a repeated sequence within the partial Hebb list and directly asked if they had recognized this repeated structure. Participants gave their response by answering with "yes" or "no". Lastly, participants were asked to recall the four repeatedly presented consonants in their correct order and typed their answer into a row of four empty boxes.

3.6. Data analysis

For each experiment, we were interested in three main questions: (1) Is there a difference in learning of the repeated list structure, as reflected in immediate recall performance, between the *No Salience* and the *High Salience* condition? (2) Is there a difference in the proportion of participants who report recognition of the repeated list structure after the working memory task between the *No Salience* and the *High Salience* condition? (3) Is there a difference in participants' ability to recall the repeated list structure after the experiment between the *No Salience* and the *High Salience* condition? We describe the analytical models used to analyze these questions in detail below.

All analyses were conducted in a Bayesian framework, and we estimated Bayes Factors to quantify the evidence in favor or against a difference between the two salience conditions for the three measures of interest. We estimated Bayes Factors using the Savage-Dickey Density Ratio (Wagenmakers et al., 2010), which compares the probability of a specific parameter value under the prior distribution to the probability of the same parameter value under the posterior distribution. In our case, all analytical models included a parameter reflecting the difference

[2] The timing in our study is slightly different from the timing in Hitch et al. (2005), who presented each letter for 500 ms with an inter-stimulus-interval (ISI) of 500 ms. We shortened the ISI to reduce the baseline performance in Filler trials, which increases the sensitivity for observing performance improvements due to sequence repetitions in the partial Hebb trials.

between the *No Salience* and the *High Salience* condition. Estimating Bayes Factors for this parameter value being equal to 0 provides evidence in favor or against the hypothesis of a difference between the two conditions. We used Cauchy priors with location = 0 on all effect parameters of our models.[3] We varied the scale of the prior with values of 0.5, 0.75, 1, and 1.25 to assure robustness of the obtained results to variations of the prior. Additionally, all Bayes Factors were re-estimated five times to ensure stability of the estimated results. In our results section, we report the median Bayes Factor over all prior scales and re-estimations, together with the obtained range.

3.6.1. Analysis of the learning effect in the working memory task

Learning during the working memory task can generally be defined as an increase in the probability of giving a correct response in the partial Hebb list compared to Filler lists over repetitions of the partial Hebb list. To analyze this, we used a Bayesian hierarchical logistic regression model with a binomial likelihood, and modeled the probability of a correct response (θ) given the number of correctly recalled items within each trial (n out of k correct) following Eq. 1:

$$\theta = logit^{-1}\big(\beta_0 + \beta_1{*}repetition + \beta_2{*}salience + \beta_3{*}repetition{*}salience + \beta_4{*}repetition{*}trialType + \beta_5{*}repetition{*}trialType{*}salience\big)$$
$$n \sim Binomial(k, \theta) \tag{1}$$

Here, the *repetition* variable reflects the number of previous presentations of the partial Hebb list and was entered into the model as a continuous predictor, starting at the value 0 and scaled to a standard deviation of 0.5 (Gelman et al., 2008). The *trialType* variable reflects if the current trial was a Hebb- or Filler-list and was dummy-coded with Filler lists = 0 and Hebb lists = 1. The *salience* variable reflects the assignment to one of the two salience groups and was effect coded with *No Salience* = −0.5 and *High Salience* = 0.5. In this specification of the model, the interaction between *repetition* and *trialType* reflects the increase in the probability of giving a correct response on the partial Hebb list compared to Filler lists over repetitions, hence, the learning effect of interest. The three-way interaction between *repetition, trialType* and *salience* reflects the difference in the learning effect between the two salience conditions. In the result section, we focus on reporting the evidence in favor or against a difference in the learning effect between the two salience groups.

The described analytical model includes the assumption that learning is reflected in a gradual increase in performance over repetitions which starts with the first presentation of the repeated Hebb list. Generally, this approach can provide an appropriate test to what extent the overall learning effect is affected by the salience manipulation. However, as has been shown by Musfeld et al. (2023a), this assumption is often an oversimplification of the underlying learning process and does not appropriately reflect the learning process on the level of individual participants. The learning curves of individuals are often rather steep but vary in when the learning process begins. To account for this, Musfeld et al. (2023a) introduced a Bayesian hierarchical mixture modeling approach, which includes additional parameters to describe the learning effect on the level of individuals: (1) a mixture proportion, which describes the proportion of participants in a sample who have shown evidence of learning; (2) a parameter for the onset point of the learning effect; (3) a parameter for the rate of the learning process. The exact specification of the model can be found in Musfeld et al. (2023a). Using this model, the authors have shown that the onset of the individual learning effects was closely tied to the timepoint in the experiment at which participants recognized the repetition. This means that differences in the mixture proportion and in the onset point of learning seem to be related to differences in the ability to recognize a repeated pattern, whereas differences in the learning rate parameter can be associated with differences in the learning process itself.

Here, we applied this model separately to each between-subject condition of the six experiments and compared the estimated posterior distributions for all parameters of interest between the two salience conditions. We hypothesized that the salience manipulation should only affect the mixture proportion and the onset point of the learning process, but not the rate of the learning effect. However, applied to our design, the model could only provide rather unprecise estimates of the described parameters. This is because participants were only able to learn parts of the Hebb list, rendering individual learning curves much noisier compared to a classic Hebb experiment in which the whole list can be learned. Thus, we only consider the results of the model to show an overall tendency in our data, but do not draw statistical inference from them.

In the analyses reported here, we included the data from all list positions, that is, repeated and unrepeated list positions. We also preregistered to conduct the same analyses by only considering those list positions which contained the repeated items in the corresponding experiments. These analyses led to the same conclusions, and we report their results in the supplementary materials (see Table S2 and Figs. S3 and S4).

3.6.2. Analysis of recognition of repetition

To analyze the difference in the proportion of participants who reported to have recognized the repeated list structure between the two salience conditions, we modeled the probability of recognizing the repetition (θ) as a function of the salience condition, using a simple Bayesian logistic regression model, with n for the number of participants in a group who recognized the repetition, and k for the group size

$$\theta = logit^{-1}(\beta_0 + \beta_1{*}salience)$$
$$n \sim Binomial(k, \theta) \tag{2}$$

The *salience* variable was again effect coded with *No Salience* = −0.5 and *High Salience* = 0.5, and therefore, the parameter estimate reflects the difference in the probability of recognizing the repetition between the two salience conditions.

3.6.3. Analysis of long-term memory recall

To analyze differences in participants' ability to recall the repeated list items in their correct order at the end of the experiment, we again modeled the probability of giving a correct response as a function of the salience condition by a Bayesian logistic regression model, using the same model equation as presented in Eq. 2. Here, the *salience* variable reflects the difference in the probability of giving a correct response in the final recall task between the two salience conditions.

4. Experiments 1–3: Does repetition learning depend on repetitions at the beginning of a list?

The findings by Hitch et al. (2005) led to the conclusion that repetitions have to be presented at the beginning of a list for sequence learning to occur. This strong constraint is predicted from the assumption of a cumulative recognition mechanisms, which incrementally matches the representation of a new list to the episodic records of previously encoded lists. Mismatching items at the beginning of the list pushes the match below a threshold, and in consequence, prevents the recognition and learning of a repeating sequence (Burgess & Hitch, 2005, 2006; see also Page and Norris (2009) for a similar dependency on the beginning of a list). In Experiments 1–3, we ask if such a positional constraint is indeed a necessary assumption for models of sequence learning, or, if instead, repeated segments of a sequence can be learned

[3] In our preregistration, we specified to use Normal priors. However, we realized that this prior setting was too restrictive on the parameters of the model, which is why we deviated from this approach. The change in the prior distribution did not affect the conclusion from our results and we show the results obtained with the preregistered prior distributions in our supplementary materials (see Table S1).

more flexibly if the recognizability of the repeated information is enhanced independently of its position in the list. To test this, we conducted the *start repeat, middle repeat,* and *end repeat* experiment (Exp. 1–3, respectively), and manipulated the salience of the repeated sequence within the presented lists to increase its recognizability.

Our predictions were the following: If it is a necessary condition for sequence learning that the repeated information has to be presented at the beginning of the list, learning should occur only in the *start repeat* experiment, but not in the *middle repeat* or *end repeat* experiment. This should be independent of the salience condition. If, however, learning only depends on the recognizability of the repeated sequence within the list, regardless of its position, we should see different results depending on the list structure and salience condition: For the *start repeat* experiment, there should be learning in both salience conditions, because the repeated sequence should be easy to identify at the beginning of the list. However, for the *middle repeat* and *end repeat* experiment, we expected to see little to no learning in the *No Salience* condition, but a learning effect in the *High Salience* condition. Here, the highlighting of the repeated sequence should facilitate its recognizability and in consequence, also allow learning in these list condition.

4.1. Results

4.1.1. The effect of salience on learning

We evaluated the effect of our salience manipulation on the three measured variables of interest: the increase in performance in the working memory task over repetitions, the percentage of participants who indicated to have recognized the repetition in the partial Hebb list, and the performance on the final long-term memory recall test. Fig. 2 shows the results from the working memory task on the left, and the results of the recognition and long-term memory test on the right. Table 2 summarizes the Bayes Factors in favor of an effect of the salience manipulation for all experiments and all measures.

For the performance in the working memory task, we observed strong learning effects in both salience conditions across all three experiments. In the *start repeat* experiment, we found overwhelming evidence for learning in both the *High Salience* condition ($BF_{10} = 6.67 \times 10^9$ [$3.32 \times 10^8 – 6.18 \times 10^{12}$]), and the *No Salience* condition ($BF_{10} = 3.51 \times 10^7$ [$7.85 \times 10^4 – 2.14 \times 10^{12}$]). This was consistent with our predictions, and Bayes Factor analyses showed substantial evidence against a difference in learning between the two salience conditions (see Table 2). To our surprise, we observed similar results for the *middle repeat* and *end repeat* experiment. Here, we not only found strong evidence in favor of a learning effect in the *High Salience* condition (*middle:* $BF_{10} = 6.71 \times 10^7$ [$4.41 \times 10^6 – 4.08 \times 10^8$]; *end:* $BF_{10} = 6.01 \times 10^5$ [$8.99 \times 10^4 – 5.94 \times 10^7$]), but also in the *No Salience* condition (*middle:* $BF_{10} = 2.46 \times 10^4$ [$5.59 \times 10^3 – 1.28 \times 10^5$]; *end:* $BF_{10} = 9.25 \times 10^2$ [$4.48 \times 10^2 – 9.37 \times 10^3$]). Descriptively, the learning effects appeared slightly stronger in the *High Salience* condition compared to the *No Salience* condition, but there was no evidence in support of any difference. In the *middle repeat* experiment, evidence remained inconclusive but showed a tendency to support the absence of an effect ($BF_{10} = 0.32$ [0.23 – 0.53]). For the *end repeat experiment*, evidence remained completely inconclusive ($BF_{10} = 1.13$ [0.77 – 1.76]). Overall, these results illustrate that participants were capable of learning the partially repeating lists in all three experiments, regardless of any enhancements to the structure's salience.

Next, we turn to the results of the percentage of participants who reported recognition of the repeated sequence. Here, we observed a descriptive trend across all experiments that the repeated structure had been recognized more often in the *High Salience* condition compared to the *No Salience* condition. For the *start repeat* and *middle repeat* experiments, this was supported by moderate to strong evidence in favor of a difference in the Bayes Factor analysis. For the *end repeat* experiment, evidence remained inconclusive (Table 2).

We observed a similar pattern for performance in the final long-term memory recall task. Here, performance was better in the *High Salience* compared to the *No Salience* condition across all three experiments. For the *start repeat* and *middle repeat* experiment, this difference was confirmed by overwhelming evidence. For the *end repeat* experiment, evidence was again inconclusive.

Our results show that the salience manipulation had an effect on the final recall of the repeated sequence in all experiments. This effect was not reflected in the learning rates during the working memory task. One potential explanation for this difference might rely in the complexity of the statistical models. In the working memory task, differences in learning rates are reflected in a three-way interaction, for which much more data is needed to obtain conclusive evidence (especially in case the effects are small), as compared to the main effect in the model for the long-term memory task. To account for this, we pooled the data from the working memory task over all three experiments and again computed the Bayes Factor for the effect of the salience manipulation on the learning rates, using all available data (see Fig. 2D).[4] Although the Bayes Factor overall increased ($BF_{10} = 2.44$ [0.95 – 5.13]), the relative evidence in favor of a salience effect remained inconclusive.

4.1.2. Parameter estimates of the mixture model for learning in the WM task

Fig. 3 shows the estimated posterior distributions for the mixture proportion, the onset of learning, and the learning rate for all three experiments. Consistent with our results so far, there were only small differences in the estimated parameters between the two salience conditions. For the mixture proportion, there was a general tendency across all three experiments that the proportion of participants who showed evidence for learning was higher in the *High Salience* condition compared to the *No Salience* condition. In the *end repeat* experiment, there was also a tendency for an earlier onset of learning in the *High Salience* compared to the *No Salience* condition. However, this tendency was not present for the *middle repeat* experiment, and even turned in the opposite direction for the *start repeat* experiment. Still, the results suggest that if there was any effect of the salience manipulation, it rather affected parameters which have been related to recognizing the repetition (i.e., onset point and mixture proportion), but not so much the parameter for the learning rate itself.

4.1.3. Exploratory analysis: Learning and repetition awareness

The results of the first three experiments showed strong evidence of learning for all three experiments. However, we found no evidence that this was affected by the salience manipulation. This raises the question if participants were simply able to recognize the repeated sequence regardless of the salience manipulation, or if the recognizability of the repeated structure did not matter for learning. To further explore this, we looked into the relationship between participants' self-reported repetition awareness, and the produced learning effect during the working memory task. Although this can only provide correlative evidence on the role of recognizing the repeated sequence, it provides a check if this relationship was at least present in our data. For this, we pooled the data from all three experiments, and then split it between those participants who reported recognition of the repeated sequence, and those who did not (see Fig. 4). To compute Bayes Factors, we used the same models as described in Eqs. 1 and 2, but now used the reported repetition awareness as a between-subjects predictor, instead of the salience group. The results showed overwhelming evidence for a difference in the learning effect for both the working memory task ($BF_{10} =$

[4] An alternative approach for addressing this issue is to only pool the data from the *middle repeat* and the *end repeat* experiments, because we only expected to observe an effect of the salience manipulation in these two experiments. We report this analysis in the supplementary materials (see Figure S5 and Table S3). It yielded the same inconclusive results as the analysis reported here.

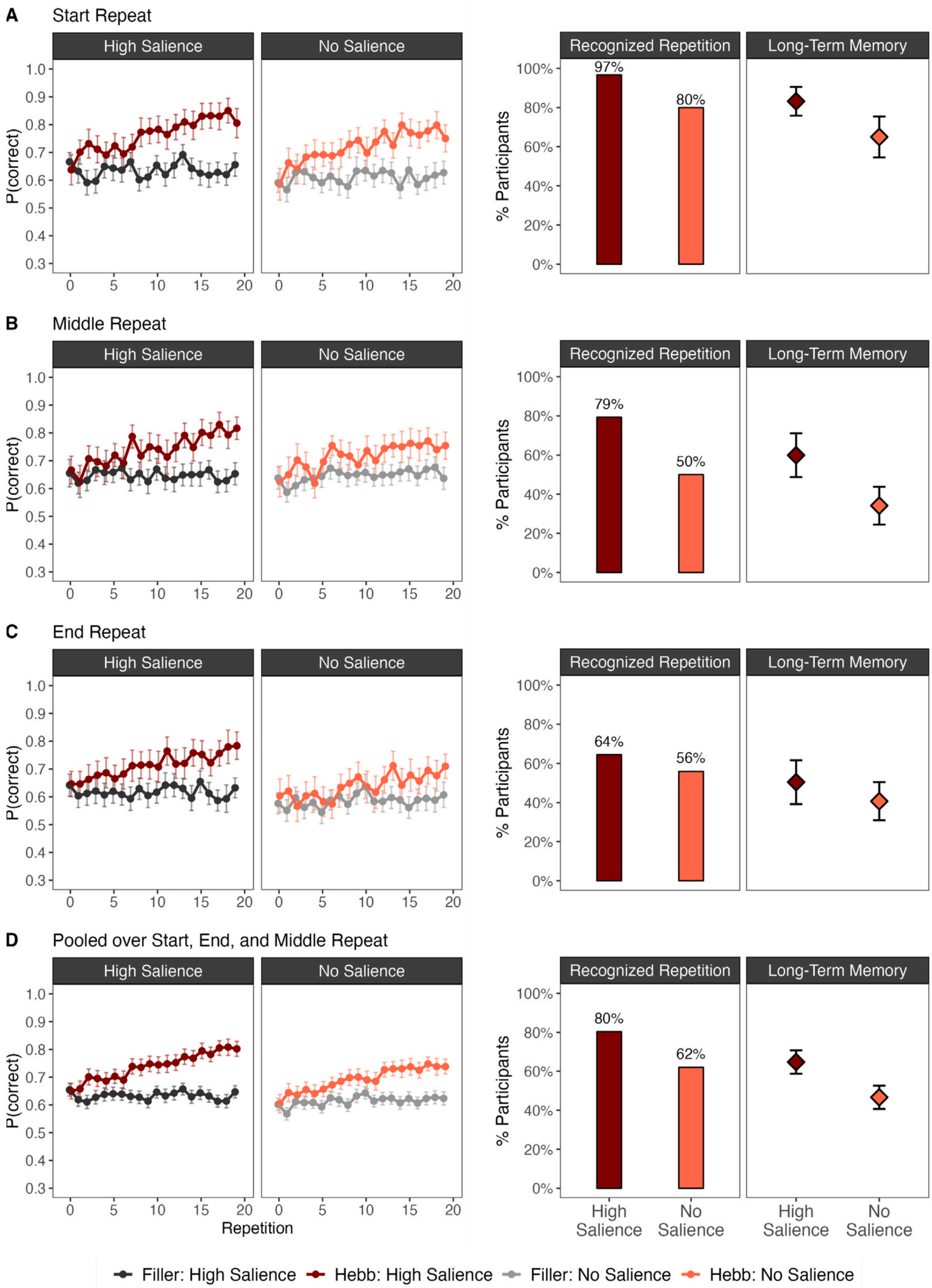

Fig. 2. Descriptive results for the start repeat (A), middle repeat (B), and end repeat (C) experiment, presenting the results of the working memory task on the left, and the results of the recognition and long-term memory recall task on the right. Panel D shows the results when data is pooled over all three Experiments. Error Bars reflect 95% within-subject confidence intervals.

Table 2
Results of the Bayes Factor analysis for the effect of salience on performance in the working memory task, percentage of participants who reported to have recognized the repetition, and performance in the final long-term memory test in Experiments 1–3. The first value shows the median, the values in parentheses the range of the obtained Bayes Factors in the prior sensitivity analysis.

	Salience Effect		
Experiment	BF_{10} ΔLearning Effect in Working Memory Task	BF_{10} ΔRecognition	BF_{10} ΔLong-Term Recall
Start Repeat	0.20 [0.14–0.34]	7.89 [5.10–9.54]	1.69×10^3 [7.57×10^2–2.67×10^3]
Middle Repeat	0.32 [0.23–0.52]	33.48 [27.55–42.86]	2.64×10^4 [8.05×10^3–4.68×10^5]
End Repeat	1.13 [0.77–1.76]	0.44 [0.33–0.59]	1.48 [0.96–1.86]
Pooled	2.44 [0.95–5.13]	125.77 [93.15–183.55]	1.74×10^6 [2.52×10^5–1.37×10^7]

7.30×10^9 [$2.22 \times 10^2 - 3.26 \times 10^{12}$]) and the long-term memory task ($BF_{10} = 1.23 \times 10^{18}$ [$9.72 \times 10^{14} - 2.68 \times 10^{21}$]), thereby confirming a strong correspondence between repetition awareness and learning.

4.2. Discussion

Experiments 1 to 3 tested if it is a necessary condition for sequence learning that the repeated information has to be presented at the beginning of the list, or if repeated information can be learned at different positions within lists, if its recognizability within the list is increased. For this, we conducted the *start repeat, middle repeat,* and *end repeat* experiment and manipulated whether the repeated structures within the lists were highlighted (*High Salience*) or not (*No Salience*). Our results provided clear evidence *against* the assumption that repetitions have to be presented at the beginning of the list. Instead, participants learned the repeated sequence at all three positions within the list. This is inconsistent with previous findings by Hitch et al. (2005) and the assumption of a cumulative recognition mechanism, which incrementally matches the episodic record of previous list encounters from the beginning of the list. Instead, it shows that people are much more flexible in learning repeated patterns than previously assumed. Against our hypotheses, the observation of learning in the *middle repeat* and *end repeat* experiment was not limited to the *High Salience* condition. Here, we observed equally strong learning effects also in the *No Salience* condition, suggesting that the salience manipulation did not affect learning.

The absence of an effect of the salience manipulation in Experiments 1–3 calls our assumption into question that repetition learning effects can be facilitated, when the recognizability of the repeated information within the list is increased. We see two possible reasons why this was the case: First, our assumption could be wrong, and the recognizability of the repeated information within the list does not matter for repetition learning. Although this appears rather unlikely considering the strong relationship we observed in our exploratory analysis between learning and repetition awareness, it cannot be ruled out by our results. The second possibility is that people might have been able to easily recognize the repetition in all list structures, even if there was no additional highlighting. In this case, we reached a ceiling effect, and the salience manipulation could not contribute much to further facilitating recognition and subsequent learning of the repeated sequence. This possibility is consistent with the rather high number of participants who reported to have recognized the repeated sequence even in the *No salience* conditions (50–80%), but it leads to the question why participants were able to easily recognize the repetitions in the *middle repeat* and *end repeat* condition in our study, but not in the study by Hitch et al. (2005).

There are three important differences between our experiments and the ones by Hitch et al. (2005), which might account for such a difference. First, we spatially arranged the presented items side-by-side in both conditions, whereas Hitch et al. (2005) presented all items in the center of the screen. An arrangement of items to distinct positions in space has not only been shown to improve short-term retention of items (e.g., Yousif et al., 2021), but also long-term associations in sequence learning (e.g., Darling et al., 2020). Thus, the spatial arrangement in the *No Salience* condition could have facilitated the recognizability of the repeated pattern within the lists even without any additional highlighting, leading to a ceiling effect in pattern recognizability.

A second difference to Hitch et al. (2005) is the number of repetitions: we repeated the partial Hebb list 20 times throughout the experiment, whereas Hitch et al. (2005) only showed four repetitions of the same partial sequence. To account for this difference, we re-analyzed the data from Experiments 1–3 and estimated the evidence for learning effects in the two salience conditions, when only considering the first four presentations of the partial Hebb list. The results of this analysis are presented in Table 3. We obtained evidence for credible learning effects in only two out of the six conditions. For all other conditions, evidence was inconclusive or even supported the absence of a learning effect. When comparing these results to the study by Hitch et al. (2005), it should be noted that they are not directly comparable, because Hitch et al. (2005) did not interleave the partially repeated sequences with additional Filler trials. Yet, it suggests that that participants in Hitch et al. (2005) might not have had enough repetitions and thereby exposure to the repeated lists, in order for stable learning effects to occur.

The last difference between the two studies is that Hitch et al. (2005) varied the set size between consecutive repetitions, whereas set size was fixed in our experiments. This has an effect on the absolute position of the repeated sequence within the lists. In our case, repetitions in the *start repeat* condition always appeared at positions 1, 2, 3, and 4, and repetitions in the *end repeat* condition always appeared at positions 5, 6, 7, and 8. In the study by Hitch et al. (2005), the absolute position of the repeated sequence was only fixed for the *start repeat* condition, but varied in the *end repeat* condition with different set sizes. For example, for a list with set size 10 in the *end repeat* condition, the repeated items were presented at positions 7, 8, 9, and 10, but for a list with set size 12, the repeated items moved to positions 9, 10, 11, and 12.

All these differences could have facilitated the ease of recognizing the repeated sequence within the list in our study, thereby leading to learning regardless of any salience manipulations. We address this potential ceiling effect in Experiment 4 by further decreasing the recognizability of the repeated sequence within the partial Hebb list.

5. Experiment 4: Can learning of repeated sequences be facilitated by increasing their recognizability?

Experiment 4 aimed to shed light on why our salience manipulation had so little effect in Experiments 1–3. If this arose because it was relatively easy to identify the repeating sequence in all three list structures, the effects should become stronger in a list condition in which it is harder to identify the repeating segment without further highlighting. To test this, we implemented another list structure, which was similar to the *chaining* condition used by Hitch et al. (2005). In this condition, we again repeated a coherent string of four consonants within the partial Hebb list, but this time, this repeated string was not presented at a fixed

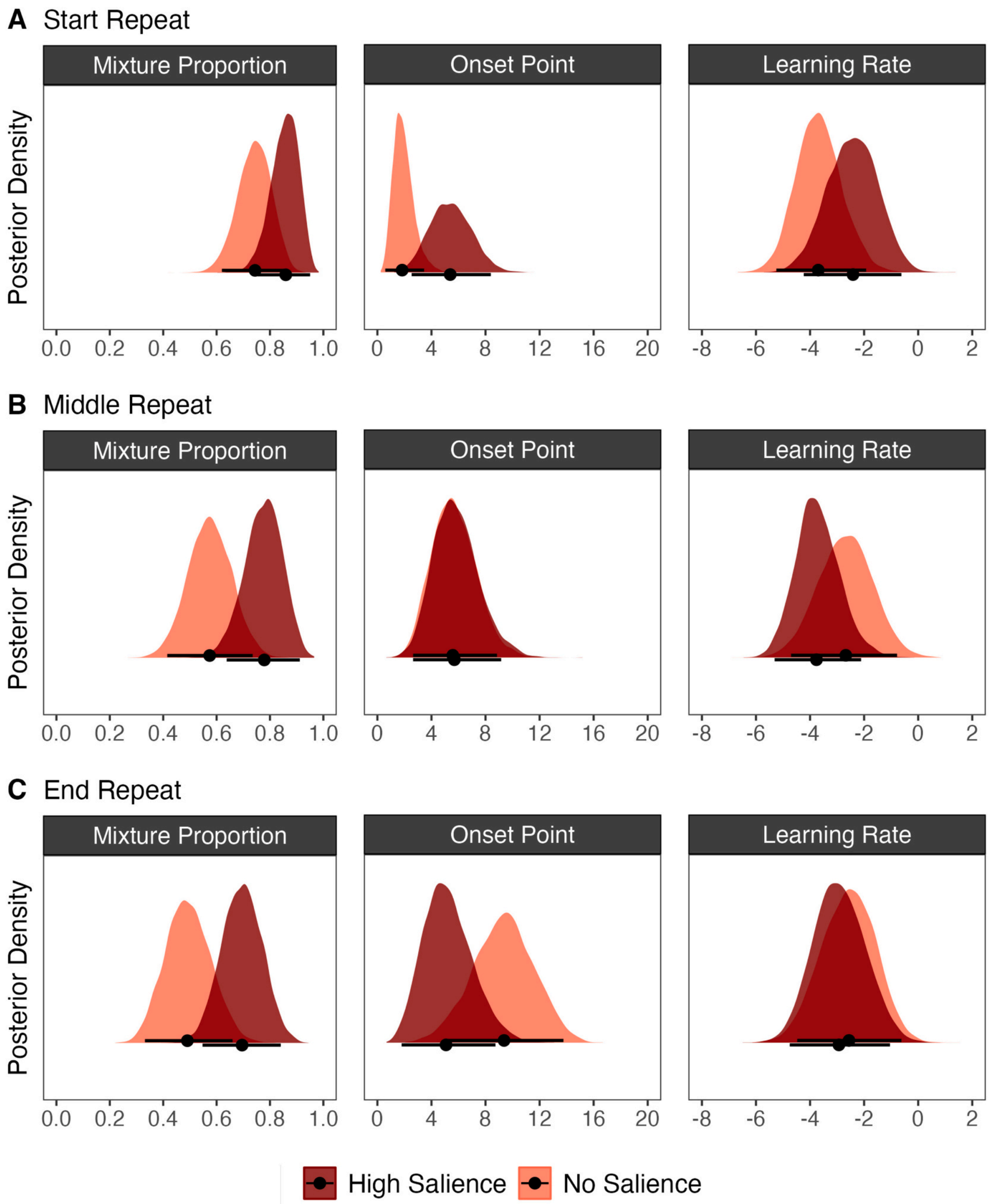

Fig. 3. Estimated posterior distributions of the mixture proportion, onset point and rate of learning in the high salience and no salience condition for the first three experiments. Points display the median of the posterior distribution. Black bars reflect the 95% Highest Density Interval.
Note. The Learning Rate is estimated on the logit scale, which is why negative values are possible. The absolute values have no direct interpretation and can only be interpreted by comparison. Larger values reflect higher learning rates.

position within the list but changed position across repetitions. This should make it much harder to identify the repeating sequence across repetitions if it is not highlighted (see Fig. 1 for a visualization of the highlighting scheme). Our prediction was the following: If repetition learning depends on recognizing the repeated sequence within the list, learning should be facilitated in the *High Salience* compared to the *No Salience* condition. In the *No Salience* group, it should be very difficult to identify the repeating string of letters, therefore preventing any learning effects. In the *High Salience* condition, however, the highlighting of the position of the repeated sequence within the list should increase its recognizability and thus, eventually also facilitate its learning. If, instead, recognizability of the repeated information does not matter for sequence learning, we would expect to see no difference between the two salience conditions. This would challenge our assumption about the importance of repetition recognizability for sequence learning and require a reconsideration of the proposed mechanisms of Musfeld et al. (2023a).

5.1. Results

5.1.1. The effect of salience on learning

The results are presented in Fig. 5, with the data from the working memory task on the left, and the data from the recognition and long-term memory task on the right. The Bayes Factors in favor of an effect of the salience manipulation are presented in Table 4.

We again observed overall credible learning effects in both, the *High*

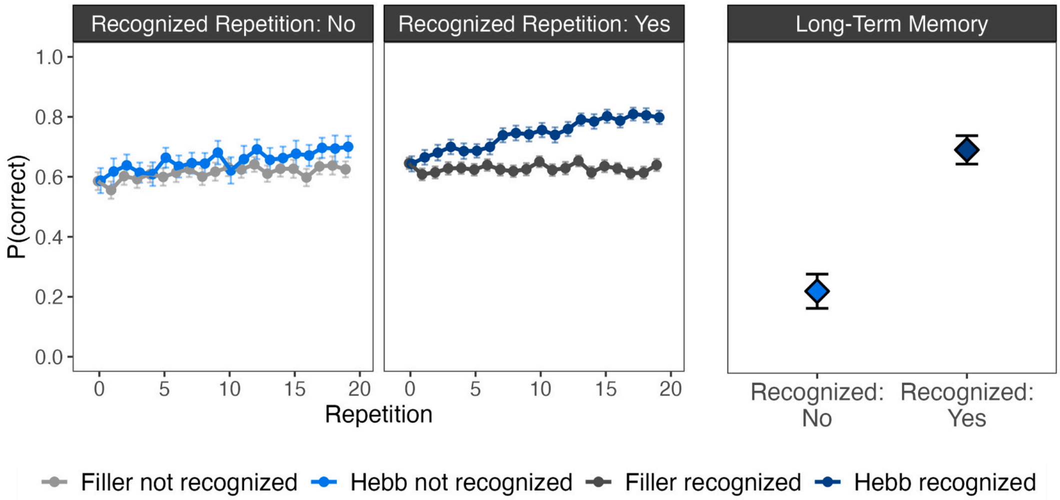

Fig. 4. Results of the Working Memory Task and Long-Term Memory task pooled over the data of Experiment 1–3, and split into participants who indicated to have recognized the repeated sequence and those who indicated to have not recognized the repeated sequence. Error Bars reflect 95% within-subject confidence intervals.

Table 3
Results of the Bayes Factor analysis for learning effects in the working memory task for both salience conditions of the start repeat, (Experiment 1), middle repeat (Experiment 2), and end repeat (Experiment3) experiment, when only considering the data from the first 4 presentations of the partial Hebb list. The first value shows the median, the values in parentheses the range of the obtained Bayes Factors in the prior sensitivity analysis.

	Evidence for Learning Effect in Working Memory Task after 4 Repetitions	
Experiment	No Salience Condition	High Salience Condition
Start Repeat	1.94 [1.42–2.97]	8.10×10^3 [2.97×10^3–2.21×10^4]
Middle Repeat	16.51 [8.62–23.83]	0.96 [0.65–1.55]
End Repeat	0.20 [0.13–0.32]	3.23 [2.22–4.72]

Salience ($BF_{10} = 1.06 \times 10^7$ [1.66×10^5–2.30×10^9]) and the *No Salience* condition ($BF_{10} = 8.25 \times 10^2$ [5.37×10^2 – 2.30×10^3]). However, this time, we found a clear difference in learning of the repeated sequence between the two salience conditions. In the *High Salience* condition, performance improvements during the working memory task were similar to Experiments 1–3, but they were strongly reduced in the *No Salience* condition. This was confirmed by the Bayes Factor analysis, showing overwhelming evidence in favor of a difference in the learning effect between the salience conditions ($BF_{10} = 760.05$ [294.45 – 1815.00]). Consistently, salience also increased the percentage of participants who reported recognition of the repeated sequence, and performance in the final long-term memory test (see Table 4).

5.1.2. Parameter estimations of the mixture model for learning in the WM task

We fitted the mixture model from Musfeld et al. (2023a) to the data of the working memory task. The results are presented in Fig. 6. As for Experiments 1–3, we found higher proportion of participants who showed evidence for learning in the *High Salience* condition compared to the *No Salience* condition (Mixture Proportion), this time accompanied by a pronounced difference in the onset point of learning. The parameter for the Learning Rate, however, showed no difference between the

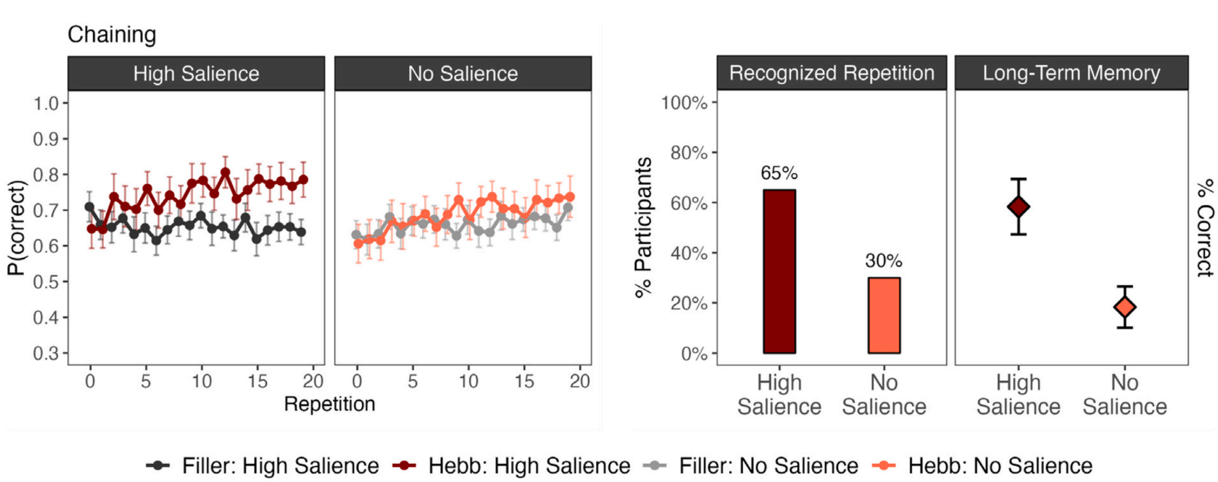

Fig. 5. Descriptive results for the chaining experiment (Experiment 4). The results of the working memory task are presented on the left, and the results of the recognition and long-term memory recall task on the right. Error Bars reflect 95% within-subject confidence intervals.

Table 4
Results of the Bayes Factor analysis for the effect of salience on performance in the working memory task, percentage of participants who reported to have recognized the repetition, and performance in the final long-term memory test in Experiment 4. The first value shows the median, the values in parentheses the range of the obtained Bayes Factors in the prior sensitivity analysis.

	Salience Effect		
Experiment	BF_{10} ΔLearning Effect in Working Memory Task	BF_{10} ΔRecognition	BF_{10} ΔLong-Term Recall
Chaining	760.05 [294.45–1815.00]	172.71 [122.51–289.66]	2.35×10^9 [1.94×10^8–3.78×10^{10}]

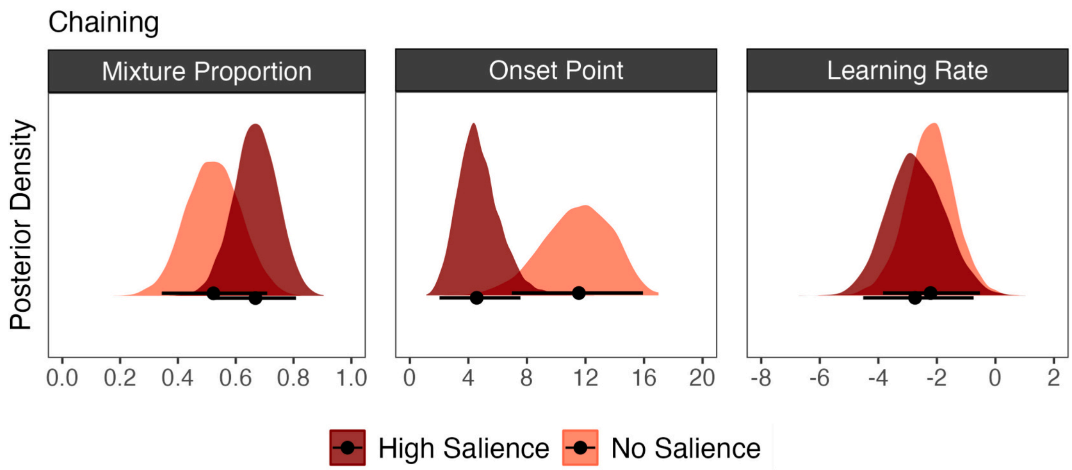

Fig. 6. Estimated Posterior Distributions of the Mixture Proportion, Onset Point and Rate of Learning in the High Salience and No Salience Condition for Experiment 4. Points Display the Median of the Posterior Distribution. Intervals reflect the 95% Highest Density Interval.
Note. The Learning Rate is estimated on the logit scale, which is why negative values are possible. The absolute values have no direct interpretation and can only be interpreted by comparison.

groups. These estimates are consistent with the assumption that the salience manipulation primarily affects participants' ability to recognize the repeated sequence within the list (as reflected by the mixture proportion and the onset point parameter), but not so much the learning process itself (as reflected by the learning rate parameter).

5.2. Discussion

We conducted Experiment 4 to test if the weak salience effects observed in Experiments 1–3 occurred because the repeated sequences were relatively easy to identify regardless of the salience condition, or if instead the identifiability of the repeated sequence within the list did not matter for learning. Under conditions that made recognizability of the repeated sequence harder (i.e., a chaining condition), we observed a strong effect of the salience manipulation on learning: In the *High Salience* condition, learning effects were comparable to those observed in Experiments 1–3. In the *No Salience* condition, learning effects were severely reduced. This shows that the recognizability of the repeated sequence within a list has a crucial role for learning, and even allows participants to learn in a chaining condition, which was not observed by Hitch et al. (2005). It further suggests that this salience effect might have been masked in the first three Experiments, because participants were able to identify the repeated sequence regardless of the salience condition, when given enough repetitions.

Taken together, the results from Experiments 1–4 provide strong evidence that learning of repeated sequences is not limited to the beginning of a list, nor to any other specific position within the list. Instead, participants learned repeated sequences even when their position changed with every repetition, as long as they were able to identify the repeated sequence within the list. This is consistent with the findings by Musfeld et al. (2023a) and emphasizes the idea that chunk formation requires people to recognize which items could be unitized into a new chunk.

6. Experiment 5–6: Do repeated items have to be presented in a coherent sequence for chunking to occur?

The results of Experiments 1–4 have shown that repetition learning can occur much more flexibly than previously assumed – that is, also when lists are only partially repeated either at the start, the middle, the end, or at varying consecutive positions. Yet, in all experiments so far, we only manipulated the position of a coherent sequence within the list, but never interrupted the coherent structure of the repeated items. Therefore, another boundary condition of the Hebb effect, apart from the recognizability of the repeated structure, might be being able to encode and output the repeated information in a coherent unit. Hitch et al. (2005) implemented list conditions in which only every other item repeated, thereby interrupting the encoding and recall of the repeated items by unrepeated information. Their results showed small benefits in recalling the items on the repeated positions, but these effects were weak, and limited to the first four positions of the list. This again raises the question if these results reflect a boundary condition of chunk formation in sequence learning, or if participants did not properly recognize the repeated structure within the list. Thus, it could be the case that chunk formation is possible even in interleaved list conditions if the recognizability of the more complex pattern is enhanced within the list.

Evidence suggesting that people might be able to learn repeated sequences even when they are interleaved with unrepeated items comes from studies which found Hebb effects in complex span tasks (Araya et al., 2022, 2023; Oberauer et al., 2015). In a complex span task, the presentation and/or recall of list items is interleaved with the presentation of distractors. Hence, in a Hebb paradigm, the presentation of the repeated list is interrupted by distractors, which could disrupt the integration of the presented items into a new chunk. Araya et al. (2022, 2023) have shown that the Hebb effect in simple and complex span tasks share the same underlying learning mechanism - which is most likely chunking. One explanation for why learning over repetitions in a

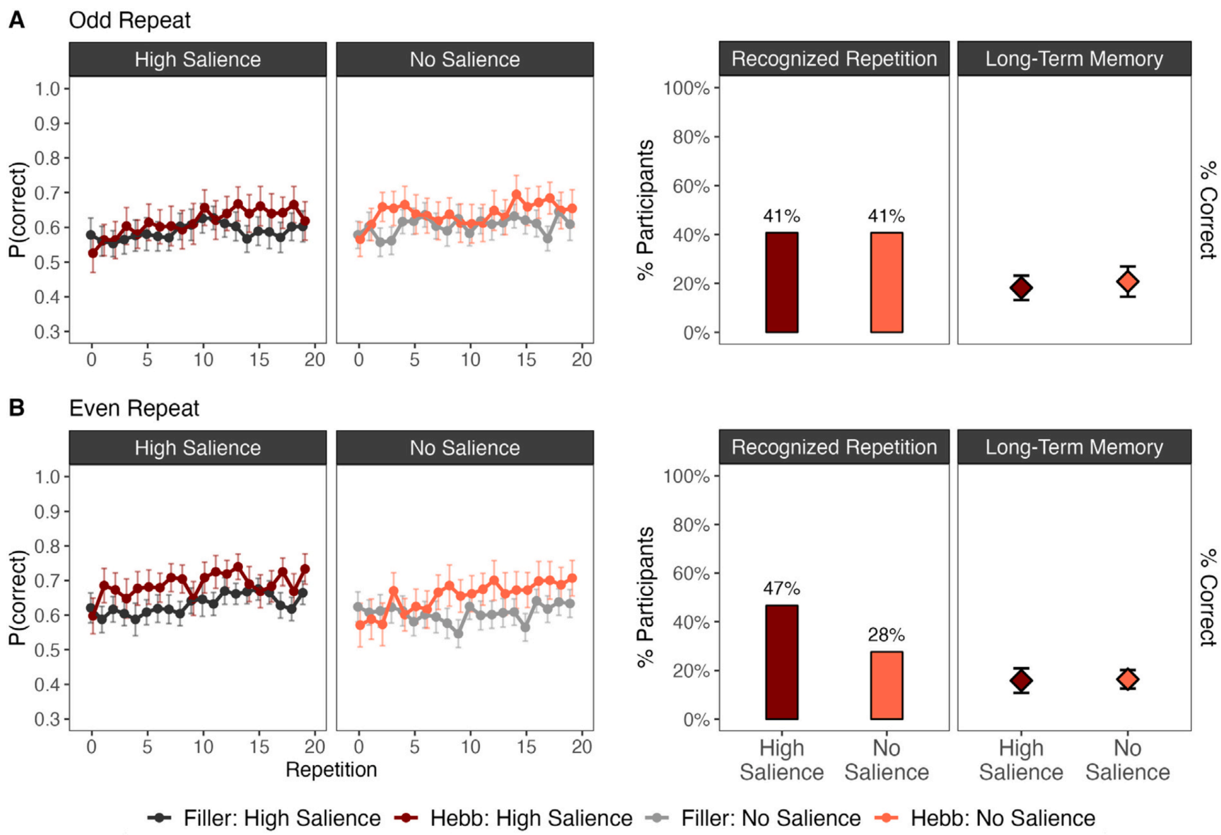

Fig. 7. Descriptive results for the odd repeat (A) and even repeat (B) experiment, presenting the results of the working memory task on the left, and the results of the recognition and long-term memory recall task on the right. Error Bars reflect 95% within-subject confidence intervals.

Table 5
Results of the Bayes Factor analysis for the effect of salience on performance in the working memory task, percentage of participants who reported to have recognized the repetition, and performance in the final long-term memory test in Experiments 5 and 6. The first value shows the median, the values in parentheses the range of the obtained Bayes Factors in the prior sensitivity analysis.

	Salience Effect		
Experiment	BF_{10} ΔLearning Effect in Working Memory Task	BF_{10} ΔRecognition	BF_{10} ΔLong-Term Recall
Odd Repeat	0.07 [0.05–0.12]	0.31 [0.22–0.46]	0.24 [0.17–0.40]
Even Repeat	0.08 [0.05–0.14]	1.98 [1.56–2.19]	0.23 [0.15–0.35]

complex span task occurs is that, in such a task, people know that the distractors are not part of the memory list, and thus, the structure of the list is very salient to them. They can then focus on integrating only the relevant items into a new chunk.

In Experiments 5 and 6, we asked if chunk formation in sequence learning can occur in conditions in which every repeated item is interleaved with an unrepeated item – in case the repeated pattern within the list is made salient. Compared to a complex span task, this is still more demanding, because participants cannot remove the interleaving unrepeated items from memory to form a coherent representation of the repeated sequence, as they still have to recall those items at test. To integrate the repeated items into a new chunk, participants might have to restructure the items during encoding into two coherent groups, and then reconstruct the sequence during recall by alternating between the groups. Alternatively, another possibility for representing and learning such a structure could be to create a template in which some of the list positions are fixed (the repeating item positions), and the other list positions are flexible and can be filled with new input during encoding. Such cognitive structures have been proposed by the Template Theory (Gobet & Simon, 1996), which has been influential to explain expertise effects in chess players. Yet, we argue that both ways of representing and learning the input most likely require participants to explicitly recognize the repeating structure within the lists.

To test if participants can form such structures, we realized two more list conditions from Hitch et al. (2005) in which either the items at the odd list positions (Experiment 5), or the items at the even list positions were repeated (Experiment 6). Again, we highlighted the repeated structure within the list in the *High Salience* group to facilitate the recognizability of the repeated items, but not in the *No Salience* group (see Fig. 1). Although being more demanding than the previously realized list conditions, we predicted that participants can eventually learn these patterns, if they can recognize them within the lists. This should be reflected in a difference in the learning effect between the *High Salience* and the *No Salience* condition because the *High Salience* condition should facilitate the recognizability of the repeated structure within the list.

6.1. Results

6.1.1. The effect of salience on learning

There was little learning in both experiments (see Fig. 7 and Table 5).

In the working memory test, performance on the interleaved Hebb lists improved only slightly above performance on Filler lists. Although Bayes Factor analyses showed that this improvement was credible in both, the *High Salience* (Odd: $BF_{10} = 29.74$ [16.31 – 51.58]; Even: $BF_{10} = 9.178 \times 10^4$ [$9.74 \times 10^3 - 2.00 \times 10^7$]) and the *No Salience* condition (Odd: $BF_{10} = 7.26$ [4.84 – 13.83]; Even: $BF_{10} = 1.05 \times 10^4$ [$2.30 \times 10^3 - 3.02 \times 10^4$]), it was much weaker than what we observed in every previous list condition. This was also reflected in the performance in the final long-term memory recall task, which was much lower than in all other experiments.

Our analysis of the effect of the salience manipulation further revealed that highlighting the repeated structure within the list had no effect on learning. For the working memory task, Bayes Factors showed strong evidence against a difference in learning between the two salience groups for both experiments. The same was true for the long-term memory recall task (see Table 5). For the proportion of participants who reported recognition of a repeated pattern within the lists, results slightly differed between the two experiments: In the *odd repeat* experiment, Bayes Factors again showed conclusive evidence against a difference between the two salience conditions. For the *even repeat* experiment, evidence remained inconclusive, but tended to favor the presence of a difference.

Because of the overall weak learning effects observed for both experiments, estimation of the parameters of the mixture model was very unprecise and did not contain any additional information. The results can be found in our supplementary material in Fig. S1.

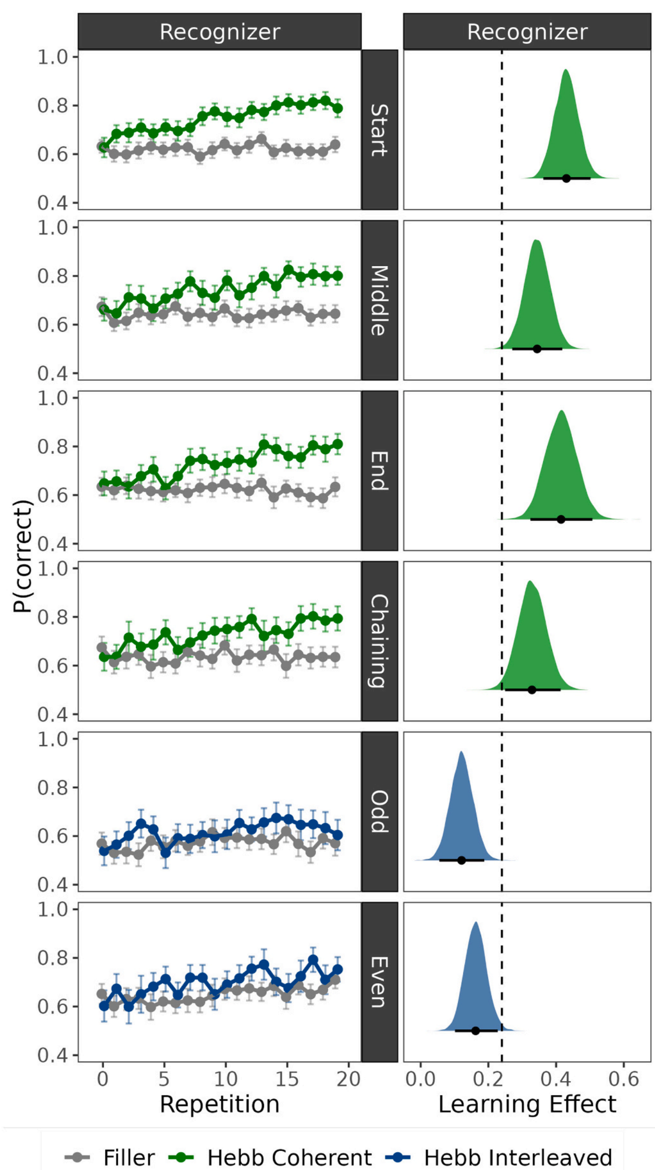

Fig. 8. Results from the working memory task of all six Experiments (left), together with the estimated learning effects (right), when only considering participants who indicated to have recognized the repeating pattern within the partial Hebb list. Error bars reflect 95% within-subject confidence intervals. Points display the median of the posterior distributions, intervals the 95% highest density interval. The dashed line provides an arbitrary visual support for emphasizing the clear difference in learning effects between the four coherent list conditions and the two interleaved list conditions.

6.1.2. *Comparison of learning effects between experiments*

The interleaved list conditions in Experiments 5 and 6 revealed overall smaller learning effects than the coherent list conditions in Experiments 1–4, suggesting that interleaved list conditions are more difficult to learn than coherent list conditions. However, in comparison to Experiments 1–4, the proportion of participants who reported recognition of the interleaved patterns was also reduced, and barely affected by our salience manipulation. Thus, it is unclear whether observed differences in the learning effect reflect differences in the ability to learn the different patterns, or whether they can be attributed to differences in the ease of recognizing these patterns.

To investigate this, we compared the learning effects between all six experiments by only considering those participants who reported recognition of the repeating pattern. This allows us to compare how well participants were able to learn the repeated pattern given that they recognized it. The results are shown in Fig. 8, with the data from the working memory task on the left, and the estimated logistic regression coefficients for the interaction between repetition and trial type (i.e., the learning effect) on the right. The estimated model coefficients can be directly compared with each other because we fitted the same model for all experiments, with the data standardized to the same scale. The results show that when controlling for differences in recognition by including only those participants who reported recognition of the repeated pattern, the learning effects were still consistently stronger in the experiments with coherent repeating sequences, compared to the two experiments with interleaved list conditions.

6.2. *Discussion*

In Experiments 5 and 6, we investigated if learning can even occur in interleaved list conditions, as long as the recognizability of the repeated pattern is increased. Our results did not support this idea. Consistent with the results reported by Hitch et al. (2005), we only found weak learning effects when repeating items in every other list position, and learning was not bolstered by highlighting the repeated pattern within the lists. Even when only considering the data of those participants who indicated recognition of the repeated pattern, learning effects were still severely reduced compared to the coherent list conditions in Experiments 1–4, showing that participants were not able to benefit from the interleaved repetitions. This adds an important boundary condition to our initial assumption: Chunk formation not only requires identifying which information could be integrated into a new chunk but also to represent this information as a coherent unit in memory.

At first glance, this conclusion might contradict chunking as an explanation for the Hebb effect in complex span tasks (Araya et al., 2022, 2023), but there is an important difference between the two tasks: In a complex span task, distractors do not have to be remembered, but can be removed from memory immediately after processing (Oberauer & Lewandowsky, 2016). Thus, they do not become part of the mental representation of the list in memory, and still allow participants to represent the list items as a contiguous sequence that can be unified into a chunk. In our task, the interleaving items could not be removed from memory, but had to be remembered at their corresponding positions

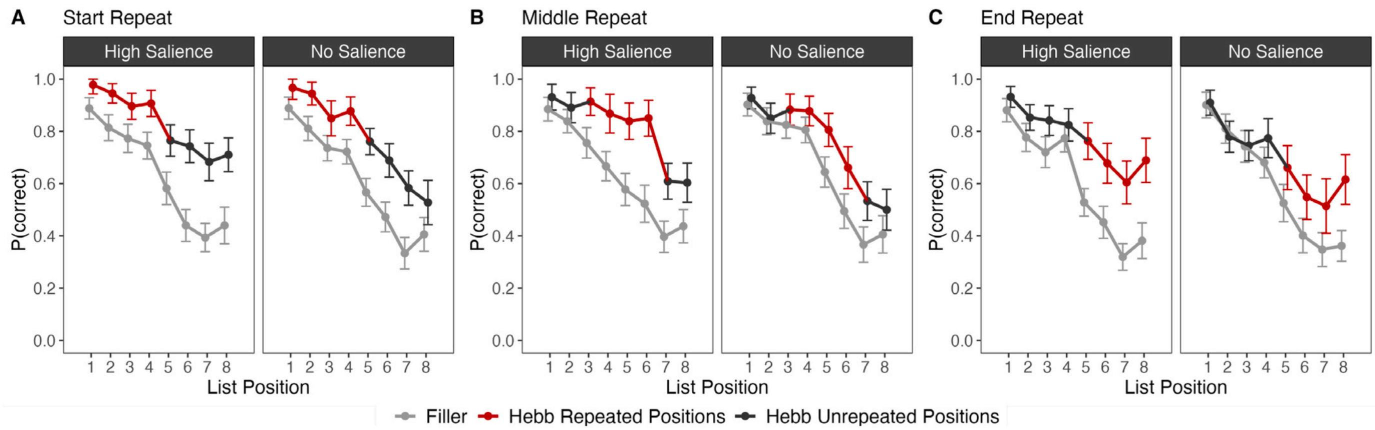

Fig. 9. Serial Position Curves for the two salience conditions in the start repeat (**A**), middle repeat (**B**), and end repeat (**C**) experiment. The presented data shows the average performance in Hebb and Filler lists from the last three mini-blocks of the working memory task as a function of list position. The positions containing the repeated sequence within the Hebb list are highlighted in red. Error Bars reflect 95% within-subject confidence intervals. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

within the list. Thus, to correctly represent the order of the whole sequence, the unrepeated items still interleaved the representation of the repeated sequence in memory and thereby prevented the integration of these items into a coherent unit.

7. Serial position curves reveal further evidence for chunk formation in repeated sequences

Previous research on chunking has shown that when pre-known chunks are combined with non-chunked items in the same lists, memory performance was not only improved for the chunk itself, but also for the other non-chunked items within the list (Bartsch & Shepherdson, 2022; Mathy et al., 2023; Mızrak & Oberauer, 2022; Norris et al., 2020; Thalmann et al., 2019). This suggests that chunks allow a more efficient representation of information in memory, thereby freeing up capacity for new information. This mostly affected items which *followed* the chunk in the list, but not so much items which *preceded* the chunk, suggesting that chunks can only free up capacity for other items proactively, but not retroactively. Our study design allows to investigate similar effects, which can provide additional evidence for a chunking mechanism in sequence learning. If the repeated sequence within the partial Hebb list was learned as a chunk, we would not only expect to see beneficial effects of repetition on the repeated list positions, but also on unrepeated list positions which follow the repeated sequence within the list. To investigate how learning of the repeated sequence within the Hebb list affected recall of items in unrepeated list positions, we plotted serial position curves for each list condition. For this, we only included the data from the last three mini-blocks of each experiment, as this is where the learning effect should be strongest and plotted the average performance in Hebb and Filler lists as a function of list positions. For the *start repeat, middle repeat,* and *end repeat* experiments, these serial position curves are presented in Fig. 9. The serial position curves for the other three experiments were less informative for analyzing pro- and retroactive effects; they can be found in Fig. S2 in the supplementary materials.

To analyze pro- and retroactive effects in the presented data, we conducted two additional analyses. For proactive effects, we combined the data from the *start repeat* and the *middle repeat* experiment, and filtered the data to those list positions, which followed the repeated sequence. Similarly, for the retroactive effects, we combined the data from the *middle repeat* and *end repeat* experiment, and filtered the data to those list positions, which preceded the repeated sequence. We then compared the average performance on the selected list positions between the partial Hebb list and the Filler lists. The results showed overwhelming evidence for improved memory performance on unrepeated list positions which followed the learned sequence ($BF_{10} = 8.20 \times 10^{12}$; proactive benefit), and strong evidence in favor of beneficial effects on list positions which preceded the learned sequence ($BF_{10} = 2242.03$; retroactive benefit). However, the proactive benefit was much larger – with an average increase of 19.47% in the probability of giving a correct response to an item on an unrepeated list position – compared to the retroactive benefit, which only led to an increase of 3.99%. This suggests that learning of the repeated sequence effectively freed up capacity for other unrepeated items in the list, as predicted from the assumption that the repeated sequence was learned as a new chunk in long-term memory.

8. General discussion

In the present study, we investigated mechanisms and boundary conditions of chunk formation in sequence learning. Chunking is a mechanism by which multiple individual elements are integrated into a single unified representation and this has been the most prominent explanation for the learning of repeated sequences (Burgess & Hitch, 2006; Mızrak & Oberauer, 2022; Page & Norris, 2009; Szmalec et al., 2009). Yet, the conditions under which chunking of repeated information occurs are not well defined.

In this study, we investigated the role of chunk recognizability in sequence learning. Recognizing what is being repeated has been shown to be a crucial condition for repetition learning to occur (Musfeld et al., 2023a; Ngiam et al., 2019). Computational models of repetition learning have incorporated the need for a recognition mechanisms but in a rather inflexible way, assuming that recognition is determined by matching repeated sequences incrementally from their beginning (Burgess & Hitch, 2006; Page & Norris, 2009). This leads to the constraint, that repeated patterns can only be recognized and learned when presented at the beginning of a sequence. Here, we argued that this assumption is too rigid, and that repeated pattern can be learned more flexibly, if their recognizability within a sequence is enhanced.

Across six experiments, we obtained strong support in favor of our proposal. Contrary to previous studies, our findings show that learning of partially repeated lists is not dependent on which part of a list is repeated. Instead, participants learned repeated sequences when presented at the beginning, in the middle, and at the end of a list - as long as they were able to recognize the repeated sequence. Even when the repeated sequence changed its position with every repetition, learning eventually occurred when the structure of the list was made salient by a spatial grouping manipulation. Only when the repeated sequence was interleaved by other unrepeated items, learning was severely reduced. These findings have important implications for our understanding of the mechanisms and boundary conditions underlying chunk formation. First, they emphasize that chunk formation is an explicit process which requires participants to recognize which items can be integrated into a new chunk. Second, our results show that the successful recognition of a repeated sequence can be guided by how information is structured during encoding. This suggests a strong link between grouping of information in working memory, and the formation of new chunks in long-term memory. In the following, we will elaborate on these conclusions in more detail, and discuss their implications for the mechanisms underlying sequence learning and the representation of order information in working memory.

8.1. Implications for mechanisms underlying sequence learning

Overall, our results are consistent with the assumption that repeated sequences are learned as chunks (Hitch et al., 2005; Mızrak & Oberauer, 2022; Page & Norris, 2009). This was not only supported by the improved memory performance observed for the repeated sequence itself, but also by the proactive (and small retroactive) effects we observed on the other items within the list. This suggests that after multiple repetitions, the repeated sequence was represented more efficiently in memory, and thereby freed up capacity for other items in the list, which is consistent with behavioral signatures of chunking (Norris et al., 2020; Thalmann et al., 2019). Apart from this, our findings also reveal new insights into mechanisms underlying chunk formation in sequence learning.

Most importantly, chunk learning in repeated sequences depends on participants' ability to recognize which items constitute a pattern of regular co-occurrence that could be learned as a new chunk. Although it has been noted before that chunking requires a marker for the onset of a chunk (see e.g., Norris & Kalm, 2021), this assumption has not yet been tested empirically. Consistent with the findings by Musfeld et al. (2023a), our results show that (1) under conditions in which recognition of the repeated pattern was easy, the learning effect was mostly driven by participants who recognized the repeated pattern (Experiments 1–3), and (2) under conditions in which it was hard to recognize the repetition, learning was facilitated by increasing the salience of the repeated pattern within the list (Experiment 4). This emphasizes that chunking is

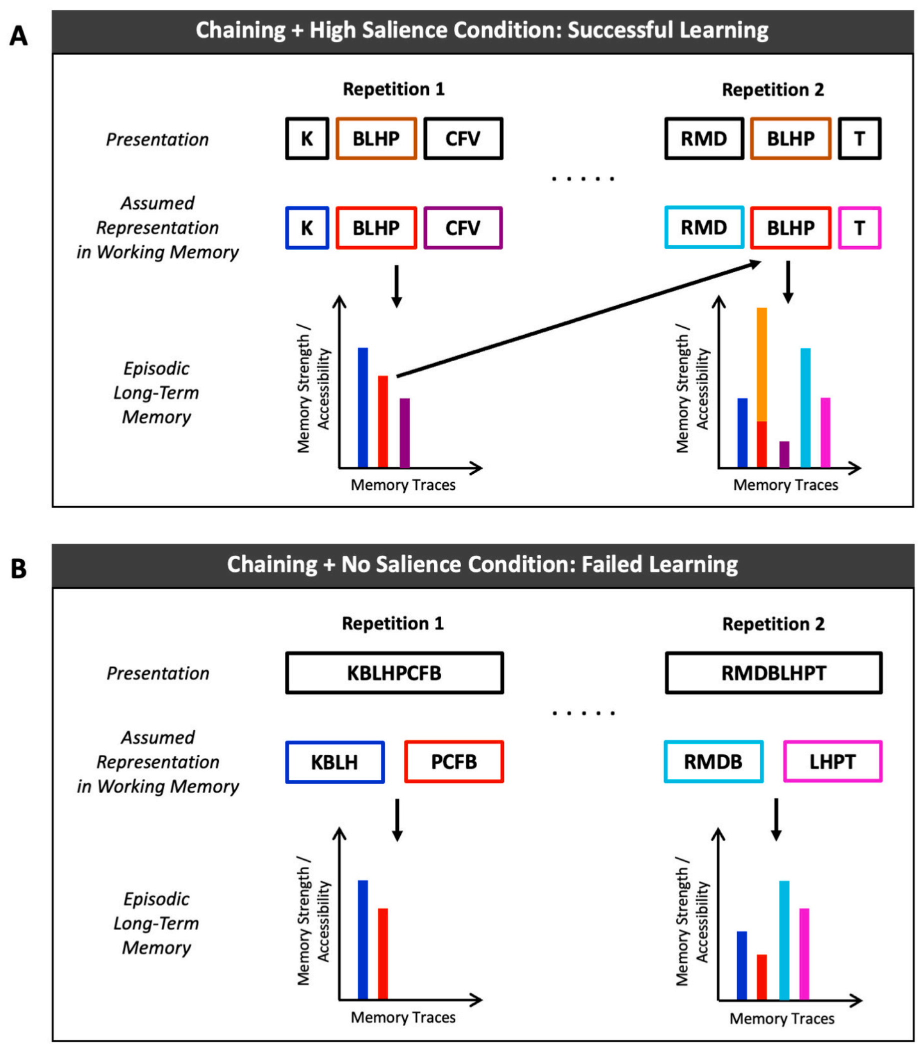

Fig. 10. Visualization of the proposed relation between the structure of representations in working memory and the learning of new chunks in long-term memory. A shows an example of the High Salience condition from the Chaining experiment (E4). The displayed structure during encoding guides how information is structured in memory, and allows an exact match of the repeated sequence, even if it appears at a different position within the list. The top row of the figure shows how the lists are presented; the middle row illustrates their grouping in working memory through frames around each group; the bottom row shows the strength of episodic-memory representations of each group, with the colors of the bars corresponding to the colors of the frames around groups in working memory. B shows an example of the No Salience condition from the Chaining experiment. No structure is given at encoding, and we assume that the list is spontaneously grouped into two groups of 4 items, shown as frames in the middle row. This leads to a mismatch between the groups encoded into episodic memory at two subsequent repetitions, and a failure in recognizing the repetition.

an explicit learning process, which does not occur implicitly with each repetition of the repeated sequence. As a consequence, models of sequence learning have to include some recognition mechanism, which determines whether the repetition of a specific sequence is explicitly recognized as having been seen before or not, before any learning can occur. The models by Burgess and Hitch (2006) and Page and Norris (2009) both already incorporate such a recognition mechanism. In their models, a representation in long-term memory is only retrieved to support recall if there is a sufficient overlap to the currently encoded list.

Yet, our results as well as the ones by Musfeld et al. (2023a) make clear that the implementations of these recognition mechanisms are both too liberal and too rigid. On the one hand, they are too liberal because they never fail in case the overlap between a presented list and a previously encoded list is sufficiently large. For a standard Hebb paradigm, in which a whole list is repeated, this always leads to successful recognition of the repeated sequence, and as a result, predicts gradual learning with every repetition. This is inconsistent with the results from the mixture model (Musfeld et al., 2023a): For many participants the onset of learning occurs only after several repetitions, and a substantial number of participants never start learning. To account for this, more elaborated recognition mechanisms are needed, which would also have to integrate assumptions on why recognizing a repeated sequence can fail, even if there is a perfect overlap with previously presented lists.

On the other hand, the recognition mechanisms in previous models (i.e. Burgess and Hitch (2006) are too rigid in that they require the current list to match a representation of a previous list from the beginning for learning to occur. In contrast to this prediction, our findings show that people were able to learn repeated sequences not only at the beginning of a list, but also in the middle or at the end of a list, and even if the position of the repeated sequence within the list changed across repetitions. This cannot be explained by the model of Burgess and Hitch (2006), and potentially causes similar problems for the model by Page and Norris (2009).

8.2. The relationship between grouping in working memory and chunk formation in long-term memory

Our findings show that current models of the Hebb effect fall short because they explain chunk formation in sequence learning on the level of an entire list. One way to overcome this would be to assume a matching and learning mechanism which does not occur on the level of a whole list, but on groups of the list. Mızrak and Oberauer (2022) came to a similar conclusion and suggested a continuous matching procedure inspired by the TRACX model (French et al., 2011), which continuously computes the overlap of the last n encoded items to chunks stored in long-term memory. However, the continuous matching mechanism implemented in the TRACX model is based on the assumption of some form of implicit sequence learning and would predict that repeated chunks can be matched and recognized regardless of their identifiability within the list. This is inconsistent with the results from our Experiment 4, because a continuous matching mechanism would predict recognition and learning of the repeated sequence regardless of the salience condition. Yet, participants only showed substantial learning effects in the *High Salience* condition of this experiment. Consequently, a continuous assessment of chunk matching seems too flexible to adequately account for our data.

As an alternative, we propose that chunk matching and subsequent learning is guided by how information is structured in working memory during encoding. We will explain this in the following (see also Fig. 10A): We first consider the *High Salience* condition in our experiments. This condition always imposes a specific structure during encoding by suggesting how to divide the presented sequence into smaller groups. We assume that each of these groups initially leaves a separate trace in episodic memory. When the partial Hebb list is presented again, the imposed grouping in the *High Salience* condition will lead participants to group the sequence in the same way as before. Each group is matched against representations in episodic memory, and if a match is detected, the repeated group is recognized as such. Thus, when partially repeated lists are consistently grouped such that the repeated part forms a separate group, the chance of recognizing the repetition is high. When recognition occurs, the episodic-memory representation found to match a current group in WM is strengthened, and that forms the basis of repetition learning. This even allows learning of the repeated sequence when it appears at different positions within the list (like in Experiment 4), as long as the repeated sequence within the list is demarcated in the displayed structure to ensure that the repeated sequence forms its own group in working memory (see Fig. 10A).

How can this account for the learning effects in the *No Salience* group? In that case, no grouping is imposed on participants. Yet, it is unlikely to assume that under this condition, participants just represent the sequence as one large unit. In fact, it has been shown that participants spontaneously divide lists into smaller groups, even when there is no prior knowledge related to these groups (Chekaf et al., 2016; Cowan & Chen, 2008; Jones & Macken, 2015; Mathy et al., 2023; Mızrak & Oberauer, 2022). Because the formation of these groups is not guided by the structure of the presented list, grouping might differ between participants, or even between trials for the same participant. For example, some participants might represent a sequence of eight letters in groups of 4–4, whereas others create groups of 3–3-2, or 2–2-2-2. This could decrease the probability of creating an exact match to a previously encoded chunk in episodic memory, and thus make learning less likely. However, if a repeated sequence is always presented at the same positions within the list, participants might eventually recognize repeating groups within the list, especially if the repeated sequence is presented at the beginning of the list. This would explain why we see substantial learning also in the *No Salience* condition of Experiments 1–3. However, when the repeated sequence moves within the list as in Experiment 4, participants' spontaneous grouping patterns will almost never match previous sub-sequences stored in episodic memory, and as a result, not lead to learning.

Our proposal suggests a strong link between grouping of encoded information in working memory, and the formation of new chunks in long-term memory. This is consistent with previous findings in the literature (Bower & Winzenz, 1969; Guitard et al., 2022; McLean & Gregg, 1967; Mızrak & Oberauer, 2022; Szmalec et al., 2009; Winzenz, 1972). For example, Bower and Winzenz (1969), and also Winzenz (1972) found that repetition learning can be impaired, when the grouping structure of a repeated list changes across repetitions. In their studies, lists of nine digits were presented in temporal (auditory) or spatial (visual) groups of one to four items. While keeping the order of items in the repeated Hebb list the same, they manipulated whether the grouping of items was the same across repetitions or changed with every repetition. Their results showed that learning only occurred when the grouping of the list remained constant. These findings on the relation between grouping and repetition learning imply that the creation of new traces in episodic long-term memory is directly guided by how information is structured in working memory during encoding.

8.3. Implications for the representation of order in working memory

The question of how the serial order of a sequence is represented in short-term and working memory has a long history, and eventually resulted in the abandonment of models relying on inter-item associations, and the acceptance of models relying on item-position associations (see Osth and Hurlstone (2023) for a recent review). Here, we don't want to revive this debate, but rather provide a clarification on how our findings fit in with the assumptions of positional models for the representation of order.

In our study, we observed strong learning effects when sequences were presented as a coherent sequence within the list (Experiments 1–4), but learning was weak when the repeated items were interleaved by other unrepeated items (Experiments 5–6). In other word, for learning to occur, a consistent association between subsequent items was more important, than a consistent association between an item and its position within the list. At first glance, this seems to contradict positional models of order in working memory. However, as Hitch et al. (2005) have already argued, the mechanisms which underly the representation of serial order in working memory don't have to be the same as the mechanisms which underly long-term learning of repeated sequences. If one assumes that these mechanisms can operate differently, our findings are in line with the assumption of positional models of order in working memory.

Specifically, considering a scenario prior to any learning effects, each item of a list can be regarded as a separate unit of information. To represent the order of a list, positional models assume that each item (and therefore each unit of information) is temporally bound to its position within the list, which then serves as a cue to retrieve the associated item during recall (Farrell, 2006; Henson, 1999; Oberauer, 2019; Oberauer et al., 2012). Moving to a scenario in which the first four items of a list repeat over multiple trials in the experiment, we assume that the repeated items are learned as a new chunk in long-term memory, which means that the separate representations of four items become integrated into one unified representation. When the repeated sequence is presented again, the chunk representation of the first four items becomes activated in long-term memory, which allows to temporally bind the whole unit to a single positional marker in working memory. This not only allows to represent the first four items of the list more efficiently, but also frees up capacity in working memory because less bindings are required to represent the order of the whole list (Norris et al., 2020; Thalmann et al., 2019). This is different from a scenario in which every other item of the list is presented repeatedly. Not only does this seem to prevent the formation of an integrated representation of the repeated items, it might also not help to improve serial recall performance if such an integrated representation existed. To preserve the correct order of the presented sequence, it wouldn't be possible to bind a chunk of the repeated items to a single positional cue in working memory as this would already distort the order of the presented sequence. Instead, it would be necessary to create separate bindings between each item and its position within the list. This limits the effects of interleaved item repetitions on serial recall performance, which is consistent with our results. Taken together, the assumption that the chunking mechanism underlying sequence learning depends on coherent item-item relations between repeated items is compatible with the assumptions of positional models of WM.

8.4. Conclusion

Across six experiments we show that chunk formation in sequence learning is not dependent on the occurrence of the repeated chunk at a specific position within a list, but rather on its recognizability. This stresses the conceptualization of chunking as an explicit rather than implicit learning mechanism. Furthermore, our findings reveal a strong link between grouping in working memory and subsequent chunk formation in long-term memory: The formation of a new chunk in long-term memory is guided by how information was initially structured in working memory.

Funding

The research was funded by grants from the Swiss National Science Foundation awarded to Klaus Oberauer (grant number 100014_192204) and Lea M. Bartsch (grant number 100014_204289).

Author note

All experiments in this study were preregistered prior to data collection. Preregistrations, data, experimental code, and analysis scripts have been made publicly available on the OSF and can be accessed at https://osf.io/pb6vx/.

CRediT authorship contribution statement

Philipp Musfeld: Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Joscha Dutli:** Writing – review & editing, Validation, Methodology, Conceptualization. **Klaus Oberauer:** Writing – review & editing, Validation, Supervision, Methodology, Funding acquisition, Conceptualization. **Lea M. Bartsch:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare no conflicts of interest.

Data availability

All data/code has been made publicly available in the OSF: https://osf.io/pb6vx/.

Acknowledgements

We thank the students of the ExPra course 2023 for helping with data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2024.105795.

References

Anderson, J. R. (1993). *Rules of the mind*. Laurence Earlbaum Associates.

Araya, C., Oberauer, K., & Saito, S. (2022). The Hebb repetition effect in complex span tasks: Evidence for a shared learning mechanism with simple span tasks. *Memory & Cognition, 50*(5), 925–940. https://doi.org/10.3758/s13421-021-01261-3

Araya, C., Oberauer, K., & Saito, S. (2023). Hebb repetition effects in complex and simple span tasks are based on the same learning mechanism. *Journal of Experimental Psychology. Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0001290

Attout, L., Ordonez Magro, L., Szmalec, A., & Majerus, S. (2020). The developmental neural substrates of Hebb repetition learning and their link with reading ability. *Human Brain Mapping, 41*(14), 3956–3969. https://doi.org/10.1002/hbm.25099

Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. (2013). Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods, 45*(2), 527–535. https://doi.org/10.3758/s13428-012-0265-2

Bartsch, L. M., & Shepherdson, P. (2022). Freeing capacity in working memory (WM) through the use of long-term memory (LTM) representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(4), 465–482. https://doi.org/10.1037/xlm0001024

Bower, G. H., & Winzenz, D. (1969). Group structure, coding, and memory for digit series. *Journal of Experimental Psychology, 80*(2, Pt.2), 1–17. https://doi.org/10.1037/h0027249

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General, 138*(4), 487–502. https://doi.org/10.1037/a0016797

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, 105*(38), 14325–14329.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114*(3), 539–576. https://doi.org/10.1037/0033-295X.114.3.539

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review, 107*(1), 127–181. https://doi.org/10.1037/0033-295X.107.1.127

Burgess, N., & Hitch, G. (2005). Computational models of working memory: Putting long-term memory into context. *Trends in Cognitive Sciences, 9*(11), 535–541. https://doi.org/10.1016/j.tics.2005.09.011

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language, 55*(4), 627–652. https://doi.org/10.1016/j.jml.2006.08.005

Bürkner, P.-C. (2017). **Brms**: An *R* package for Bayesian multilevel models using *Stan*. *Journal of Statistical Software, 80*(1). https://doi.org/10.18637/jss.v080.i01

Chekaf, M., Cowan, N., & Mathy, F. (2016). Chunk formation in immediate memory and how it relates to data compression. *Cognition, 155*, 96–107. https://doi.org/10.1016/j.cognition.2016.05.024

Couture, M., & Tremblay, S. (2006). Exploring the characteristics of the visuospatial hebb repetition effect. *Memory & Cognition, 34*(8), 1720–1729. https://doi.org/10.3758/BF03195933

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review, 24*(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

Cowan, N., & Chen, Z. (2008). How chunks form in long-term memory and affect short-term memory limits. In M. P. A. Page, & A. Thorn (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 86–107). East Sussex, UK: Psychology Press.

Cumming, N., Page, M. P. A., & Norris, D. (2003). Testing a positional model of the Hebb effect. *Memory, 11*(1), 43–63. https://doi.org/10.1080/741938175

Darling, S., Havelka, J., Allen, R. J., Bunyan, E., & Flornes, L. (2020). Visuospatial bootstrapping: Spatialized displays enhance digit and nonword sequence learning. *Annals of the New York Academy of Sciences, 1477*(1), 100–112. https://doi.org/10.1111/nyas.14429

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211–245. https://doi.org/10.1037/0033-295X.102.2.211

Farrell, S. (2006). Mixed-list phonological similarity effects in delayed serial recall. *Journal of Memory and Language, 55*(4), 587–600. https://doi.org/10.1016/j.jml.2006.06.002

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review, 119*(2), 223–271. https://doi.org/10.1037/a0027371

French, R. M., Addyman, C., & Mareschal, D. (2011). TRACX: A recognition-based connectionist framework for sequence segmentation and chunk extraction. *Psychological Review, 118*(4), 614–636. https://doi.org/10.1037/a0025255

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics, 2*(4). https://doi.org/10.1214/08-AOAS191

Gobet, F., Lloyd-Kelly, M., & Lane, P. C. R. (2016). What's in a Name? The Multiple Meanings of "Chunk" and "Chunking". *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.00102

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31*(1), 1–40. https://doi.org/10.1006/cogp.1996.0011

Guérard, K., Saint-Aubin, J., Boucher, P., & Tremblay, S. (2011). The role of awareness in anticipation and recall performance in the Hebb repetition paradigm: Implications for sequence learning. *Memory & Cognition, 39*(6), 1012–1022. https://doi.org/10.3758/s13421-011-0084-1

Guitard, D., Saint-Aubin, J., & Cowan, N. (2022). Grouping effects in immediate reconstruction of order and the preconditions for long-term learning. *Quarterly Journal of Experimental Psychology, 75*(1), 70–97. https://doi.org/10.1177/17470218211030825

Hebb, D. O. (1961). Distinctive features of learning in the higher animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). Blackwell.

Henninger, F., Shevchenko, Y., Mertens, U., Kieslich, P. J., & Hilbig, B. E. (2022). *Lab.Js: A free, open, online experiment builder (v22.0-beta5) [Computer software]*. Zenodo. https://doi.org/10.5281/ZENODO.597045

Henson, R. N. A. (1999). Coding position in short-term memory. *International Journal of Psychology, 34*(5–6), 403–409. https://doi.org/10.1080/002075999399756

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *The Quarterly Journal of Experimental Psychology A, 49*(1), 80–115. https://doi.org/10.1080/027249896392810

Hitch, G. J., Fastame, M., & Flude, B. (2005). How is the serial order of a verbal sequence coded? Some comparisons between models. *Memory, 13*(3–4), 247–258. https://doi.org/10.1080/09658210344000314

Huang, L., & Awh, E. (2018). Chunking in working memory via content-free labels. *Scientific Reports, 8*(1), 23. https://doi.org/10.1038/s41598-017-18157-5

Johnson, A. J., & Miles, C. (2019). Visual Hebb repetition effects: The role of psychological distinctiveness revisited. *Frontiers in Psychology, 10*, 17. https://doi.org/10.3389/fpsyg.2019.00017

Jones, G., Gobet, F., Freudenthal, D., Watson, S. E., & Pine, J. M. (2014). Why computational models are better than verbal theories: The case of nonword repetition. *Developmental Science, 17*(2), 298–310. https://doi.org/10.1111/desc.12111

Jones, G., & Macken, B. (2015). Questioning short-term memory and its measurement: Why digit span measures long-term associative learning. *Cognition, 144*, 1–13. https://doi.org/10.1016/j.cognition.2015.07.009

Kay, M. (2023). *Tidybayes: Tidy data and Geoms for Bayesian models (v3.0.6) [Computer software]*. Zenodo. https://doi.org/10.5281/ZENODO.1308151

Lewandowsky, S., & Farrell, S. (2008). Phonological similarity in serial recall: Constraints on theories of memory. *Journal of Memory and Language, 58*(2), 429–448. https://doi.org/10.1016/j.jml.2007.01.005

Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review, 96*(1), 25–57. https://doi.org/10.1037/0033-295X.96.1.25

Logan, G. D. (2021). Serial order in perception, memory, and action. *Psychological Review, 128*(1), 1–44. https://doi.org/10.1037/rev0000253

Logan, G. D., & Cox, G. E. (2021). Serial memory: Putting chains and position codes in context. *Psychological Review, 128*(6), 1197–1205. https://doi.org/10.1037/rev0000327

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281.

Makowski, D., Ben-Shachar, M., & Lüdecke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software, 4*(40), 1541. https://doi.org/10.21105/joss.01541

Mathy, F., Friedman, O., & Gauvrit, N. (2023). Can compression take place in working memory without a central contribution of long-term memory? *Memory & Cognition*. https://doi.org/10.3758/s13421-023-01474-8

McKelvie, S. J. (1987). Learning and awareness in the Hebb digits task. *The Journal of General Psychology, 114*(1), 75–88. https://doi.org/10.1080/00221309.1987.9711057

McLean, R. S., & Gregg, L. W. (1967). Effects of induced chunking on temporal aspects of serial recitation. *Journal of Experimental Psychology, 74*(4, Pt.1), 455–459. https://doi.org/10.1037/h0024785

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97. https://doi.org/10.1037/h0043158

Mızrak, E., & Oberauer, K. (2022). Working memory recruits long-term memory when it is beneficial: Evidence from the Hebb effect. *Journal of Experimental Psychology: General, 151*(4), 763–780. https://doi.org/10.1037/xge0000934

Müller, K. (2020). here: A simpler way to find your files (1.0.1) [computer software]. https://CRAN.R-project.org/package=here.

Murdock, B. B. (1995). Developing TODAM: Three models for serial-order information. *Memory & Cognition, 23*(5), 631–645. https://doi.org/10.3758/BF03197264

Musfeld, P., Dutli, J., Oberauer, K., & Bartsch, L. M. (2023). *Data, Materials, and Code for: Grouping in Working Memory Guides Chunk Formation in Long-Term Memory*. https://doi.org/10.17605/OSF.IO/PB6VX [dataset].

Musfeld, P., Souza, A. S., & Oberauer, K. (2023a). Repetition learning is neither a continuous nor an implicit process. *Proceedings of the National Academy of Sciences, 120*(16), Article e2218042120. https://doi.org/10.1073/pnas.2218042120

Musfeld, P., Souza, A. S., & Oberauer, K. (2023b). Testing expectations and retrieval practice modulate repetition learning of visuospatial arrays. *Journal of Experimental Psychology. Learning, Memory, and Cognition*. https://doi.org/10.1037/xlm0001298

Ngiam, W. X. Q., Brissenden, J. A., & Awh, E. (2019). "Memory compression" effects in visual working memory are contingent on explicit long-term memory. *Journal of Experimental Psychology: General, 148*(8), 1373–1385. https://doi.org/10.1037/xge0000649

Norris, D., & Kalm, K. (2021). Chunking and data compression in verbal short-term memory. *Cognition, 208*, Article 104534. https://doi.org/10.1016/j.cognition.2020.104534

Norris, D., Kalm, K., & Hall, J. (2020). Chunking and redintegration in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(5), 872–893. https://doi.org/10.1037/xlm0000762

Norris, D., Page, M. P. A., & Hall, J. (2018). Learning nonwords: The Hebb repetition effect as a model of word learning. *Memory, 26*(6), 852–857. https://doi.org/10.1080/09658211.2017.1416639

Oberauer, K. (2009). Design for a Working Memory. In *, Vol. 51. Psychology of learning and motivation* (pp. 45–100). Elsevier. https://doi.org/10.1016/S0079-7421(09)51002-X.

Oberauer, K. (2019). Working memory capacity limits memory for bindings. *Journal of Cognition, 2*(1), 40. https://doi.org/10.5334/joc.86

Oberauer, K., Jones, T., & Lewandowsky, S. (2015). The Hebb repetition effect in simple and complex memory span. *Memory & Cognition, 43*(6), 852–865. https://doi.org/10.3758/s13421-015-0512-8

Oberauer, K., & Lewandowsky, S. (2016). Control of information in working memory: Encoding and removal of distractors in the complex-span paradigm. *Cognition, 156*, 106–128. https://doi.org/10.1016/j.cognition.2016.08.007

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., … Ward, G. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin, 144*(9), 885–958. https://doi.org/10.1037/bul0000153

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review, 19*(5), 779–819. https://doi.org/10.3758/s13423-012-0272-4

Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences, 105*(7), 2745–2750. https://doi.org/10.1073/pnas.0708424105

Osth, A. F., & Dennis, S. (2015). Prior-list intrusions in serial recall are positional. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(6), 1893–1901. https://doi.org/10.1037/xlm0000110

Osth, A. F., & Hurlstone, M. J. (2023). Do item-dependent context representations underlie serial order in cognition? Commentary on Logan (2021). *Psychological Review, 130*(2), 513–545. https://doi.org/10.1037/rev0000352

Page, M. P. A., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word forms. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 364*(1536), 3737–3753. https://doi.org/10.1098/rstb.2009.0173

R Core Team. (2023). *R: A language and environment for statistical computing (4.3.1) [Computer software]*. R Foundation for Statistical Computing. https://www.R-project.org/.

Robinet, V., Lemaire, B., & Gordon, M. B. (2011). MDLChunker: A MDL-based cognitive model of inductive learning. *Cognitive Science, 35*(7), 1352–1389. https://doi.org/10.1111/j.1551-6709.2011.01188.x

Saint-Aubin, J., & Guérard, K. (2018). The Hebb repetition effect as a laboratory analogue of language acquisition: Learning three lists at no cost. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale, 72*(1), 2–8. https://doi.org/10.1037/cep0000136

Schwartz, M., & Bryden, M. P. (1971). Coding factors in the learning of repeated digit sequences. *Journal of Experimental Psychology, 87*(3), 331–334. https://doi.org/10.1037/h0030552

Solway, A., Murdock, B. B., & Kahana, M. J. (2012). Positional and temporal clustering in serial order memory. *Memory & Cognition, 40*(2), 177–190. https://doi.org/10.3758/s13421-011-0142-8

Souza, A. S., & Oberauer, K. (2022). Promoting visual long-term memories: When do we learn from repetitions of visuospatial arrays? *Journal of Experimental Psychology. General*. https://doi.org/10.1037/xge0001236

Stan Development Team. (2023). RStan: The R interface to Stan (2.26.23) [Computer software]. https://mc-stan.org/.

Sukegawa, M., Ueda, Y., & Saito, S. (2019). The effects of Hebb repetition learning and temporal grouping in immediate serial recall of spatial location. *Memory & Cognition, 47*(4), 643–657. https://doi.org/10.3758/s13421-019-00921-9

Szmalec, A., Duyck, W., Vandierendonck, A., Mata, A. B., & Page, M. P. A. (2009). Short article: The Hebb repetition effect as a laboratory analogue of novel word learning. *Quarterly Journal of Experimental Psychology, 62*(3), 435–443. https://doi.org/10.1080/17470210802386375

Thalmann, M., Souza, A. S., & Oberauer, K. (2019). How does chunking help working memory? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(1), 37–55. https://doi.org/10.1037/xlm0000578

Tulving, E. (1972). Episodic and semantic memory. In E. Tulving, & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). Academic Press.

Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. *Cognitive Psychology, 60*(3), 158–189. https://doi.org/10.1016/j.cogpsych.2009.12.001

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org/10.21105/joss.01686

Winzenz, D. (1972). Group structure and coding in serial learning. *Journal of Experimental Psychology, 92*(1), 8–19. https://doi.org/10.1037/h0032161

Yousif, S. R., Rosenberg, M. D., & Keil, F. C. (2021). Using space to remember: Short-term spatial structure spontaneously improves working memory. *Cognition, 214*, Article 104748. https://doi.org/10.1016/j.cognition.2021.104748