A comparison of serial order short-term memory effects across verbal and musical domains

Simon Gorin[a,b], Pierre Mengal[a], and Steve Majerus[a,b]

[a] Psychology and Neuroscience of Cognition Research Unit (PsyNCog), University of Liège, Liège,

Belgium

[b] Fund for Scientific Research–FNRS, Brussels, Belgium

Author note

Correspondence concerning this article should be addressed to Simon Gorin, who is now at

the Department of Cognitive Development, Université de Genève, 40 Boulevard du Pont d'Arve,

1205 Genève, Switzerland; e-mail: gorinsimon@gmail.com; phone: +41223799278.

All relevant data are available through the Open Science Framework (https://osf.io/6kvrz/).

Abstract

Recent studies suggest that the mechanisms involved in the short-term retention of serial order information may be shared across short-term memory (STM) domains such as verbal and visuospatial STM. Given the intrinsic sequential organization of musical material, the study of STM for musical information may be particularly informative about serial order retention processes and their domain-generality. The present experiment examined serial order STM for verbal and musical sequences in participants with no advanced musical expertise and experienced musicians. Serial order STM for verbal information was assessed via a serial order reconstruction task for digit sequences. In the musical domain, serial order STM was assessed using a novel melodic sequence reconstruction task maximizing the retention of tone order information. We observed that performance for the verbal and musical tasks was characterized by sequence length as well as primacy and recency effects. Serial order errors in both tasks were characterized by similar transposition gradients and ratios of fill-in:infill errors. These effects were observed for both participant groups, although the transposition gradients and ratios of fill-in/infill errors showed additional specificities for musician participants in the musical task. The data support domain-general serial order STM effects but also suggest the existence of additional, domain-specific effects. Implications for models of serial order STM in verbal and musical domains are discussed.

Keywords: serial order, working memory, music cognition, musical expertise, language and music

Recent theoretical accounts of serial order STM have proposed the existence of domain-general serial order STM mechanisms. This is supported by studies comparing verbal and visuospatial STM domains, showing similar hallmark serial order STM effects across the two domains (for a recent review, see Hurlstone, Hitch, & Baddeley, 2014). The aim of this study was to provide further evidence for the domain-generality hypothesis of serial order mechanisms in STM, by examining serial order phenomena across verbal and musical domains of STM. Given the intrinsic temporal and sequential nature of music, the study of STM for musical material could be particularly informative regarding the nature of serial order processes in STM and their domain-generality.

In the verbal domain, STM for serial order information is typically assessed by requiring participants to recall in forward serial order sequences of familiar information such as digits, letters, or words. This type of task has led to well-replicated benchmark effects of serial order in STM, such as specific serial position curves for response accuracy and transposition gradients for erroneous responses (for a review, see Hurlstone et al., 2014). As regards response accuracy, a U-shaped serial recall curve is typically observed, with marked recency and primacy effects (see, e.g., Cowan, Saults, Elliott, & Moreno, 2002; Oberauer, 2003). Concerning serial order errors, they are governed by a locality constraint (Henson, 1996), characterized by an increased proportion of transposition errors for close serial positions. Furthermore, ordering errors are characterized by the presence of approximately 2 times more fill-in than infill transpositions (see Farrell, Hurlstone, & Lewandowsky, 2013; Henson, 1996; Surprenant, Kelley, Farley, & Neath, 2005); when an item is recalled one position too soon in a sequence, that is, at position i-1, the following item is more likely to be the item actually presented at position i-1 (fill-in transposition) than the next item in the sequence (infill transposition).

These effects and others are considered as behavioral signatures and direct evidences for the involvement of specific serial order constructs (see, e.g., Farrell & Lewandowsky, 2004; Hurlstone et al., 2014; Lewandowsky & Farrell, 2008), and are at the heart of several computational models of serial order in the verbal domain of STM (Botvinick & Plaut, 2006; Brown, Neath, & Chater, 2007; Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1992, 1999, 2006; Farrell & Lewandowsky, 2002; Henson, 1998; Lewandowsky & Farrell, 2008; Page & Norris, 1998, 2009). Most of these models assume that the identity of items presented in a list of memoranda, and their order of occurrence

inside the list, are represented by distinct mechanisms (but see Botvinick & Plaut, 2006). The maintenance of item information is supported by item-specific representations that, in some models, reflect temporary activation of linguistic long-term memory, while serial order information is considered to be represented through specific context signals that track the occurrence of items in a memory list. According to ordinal models of serial order, the context signal takes the form of a primacy activation gradient, which allows the encoding of successive items with decreasing strength (e.g., Farrell & Lewandowsky, 2002; Page & Norris, 1998). In positional models, serial order information is represented through the association between items and the different states of a dynamic contextual signal; the contextual signal has been proposed to reflect timing-based or episodic-based representations (Brown et al., 2000; Burgess & Hitch, 2006; Henson, 1998; Lewandowsky & Farrell, 2008). The different categories of models mentioned here are not only able to account for all major serial recall phenomena, but it has been recently proposed that the principles underlying the representation of serial order implemented in these models may account for serial recall phenomena across different domains, such as verbal and visuospatial STM modalities (for a review, see Hurlstone et al., 2014). It is also important to emphasize that the view considering the existence of amodal serial order coding mechanisms in STM is not a recent one (Jones, Farrand, Stuart, & Morris, 1995; Jones & Macken, 1993; Ward, Avons, & Melling, 2005).

The suggestion of similar serial order coding mechanisms across verbal and visuospatial STM domains stems from the observation of similar serial order phenomena in verbal and visuospatial STM tasks. For example, in the verbal domain, the serial response curve adopts a characteristic U-shape, with marked recency and primacy effects. This U-shaped response curve has also been observed for recall accuracy of lists of visuospatial memoranda, such as faces, spatial locations or visual configurations (Avons, 1998; Guérard & Tremblay, 2008; Jones et al., 1995; Smyth, Hay, Hitch, & Horton, 2005). There is also a close resemblance of serial order error patterns in verbal and visuospatial STM tasks: In both domains, errors are characterized by transposition gradients governed by a locality constraint (e.g., Henson, 1996; Parmentier, Andrés, Elford, & Jones, 2006; Parmentier, King, & Dennis, 2006; Smyth et al., 2005), and by more fill-in than infill errors (Farrell et al., 2013; Guérard & Tremblay, 2008; Henson, 1996; Surprenant et al., 2005). Finally, the effect of temporal

grouping (see, e.g., Hartley, Hurlstone, & Hitch, 2016; Ryan, 1969) has also been observed in both verbal and visuospatial domains (Hurlstone & Hitch, 2015; Parmentier, Andrés, et al., 2006; Parmentier, Maybery, & Jones, 2004). The temporal grouping effect arises when groups of items in a stimulus sequence are marked via the insertion of temporal pauses between the last item of a group and the first item of the next group, leading to higher recall performance for grouped versus ungrouped lists of items. These cross-domain commonalities of serial order phenomena have led to the hypothesis that serial order STM processes may, at least partially, be domain-general (Hurlstone & Hitch, 2015, 2017).

Compared to verbal and visuospatial domains of STM, the musical STM domain has been poorly studied, particularly as concerns the existence of the hallmark serial order phenomena discussed above. Yet STM for music may be a heuristically important domain for exploring serial order STM processes, given the intrinsically sequential nature of musical information, which, at the most basic level, is characterized by sequences of tones. Some studies suggest the possible existence in musical STM of similar serial order phenomena as observed for verbal and visuospatial STM. Primacy and recency effects have been observed for STM recall or recognition of tone sequences (Gorin, Mengal, & Majerus, 2017; Greene & Samuel, 1986; Mondor & Morin, 2004). Also, there is some indication that in musical reproduction tasks, serial order transpositions may follow similar transposition gradients as is observed in verbal and visuospatial STM tasks (Mathias, Pfordresher, & Palmer, 2015). However, it should be noted that these musical production tasks differed from STM tasks in several ways, such as involving a learning phase of the to-be-produced sequences. More generally, while models of musical production acknowledge a critical role for serial order STM in producing musical sequences (see, e.g., Pfordresher, Palmer, & Jungers, 2007), models of musical STM do currently not directly address the problem of serial order. In his model, Berz (1995) proposed to add a musical module to the multicomponent model of working memory (Baddeley & Hitch, 1974), but without considering any specific mechanisms for the representation and storage of serial order information. More recently, Ockelford (2007) suggested that a musical central executive component of STM could have the function of tracking the serial order of information via serial tagging mechanisms (see also Kieras, Meyer, Mueller, & Seymour, 1999). Overall, the musical STM literature

currently does not provide detailed accounts of the processing of serial order information. The aim of the present study is to further our understanding of serial order processing in musical STM, by considering the hypothesis that serial order processes are not specific to the musical STM domain but are shared with other domains such as the verbal STM domain.

In order to achieve this aim, we conducted a comprehensive investigation of different serial order phenomena in musical and verbal STM modalities, by focusing on four serial order hallmark effects: sequence length effect, primacy and recency effects, transposition gradients, and the ratio of fill-in and infill errors. Sequence length effect, primacy and recency effects, and transposition gradients are among the most robust and most frequently studied characteristics of performance in serial order STM tasks (see Hurlstone et al., 2014). We were also interested in the ratio between fill-in and infill errors as this effect is particularly informative as regards models of serial order STM: the presence of more fill-in than infill errors is considered as evidence against chaining accounts of serial order STM (Kieras et al., 1999; Lewandowsky & Murdock, 1989), given that chaining accounts of serial order predict more infill than fill-in errors (Henson, 1996; Surprenant et al., 2005). If similar serial order processes are involved in verbal and musical STM tasks, we should observe qualitatively similar serial order phenomena for both types of task. According to the domain-generality hypothesis, we expect to observe, in both verbal and musical tasks, (1) U-shaped serial position curves with marked primacy and recency effects, (2) a locality constraint for transposition error gradients, and (3) a 2:1 ratio between fill-in and infill errors. At the same time, it is possible that memory systems for different types of material (e.g., verbal, visual, musical) could code serial order information using similar representational properties; hence, similar behavioral patterns across modalities do not necessarily reflect the use of a common, modality-general serial order processing system (Logie, Saito, Morita, Varma, & Norris, 2016; Saito, Logie, Morita, & Law, 2008). It is important to emphasize that the aim of this study is not to determine whether musical and verbal STM systems are totally domain-specific or domain-general, but rather to assess the extent to which the maintenance of serial order information is characterized by similar effects across verbal and musical modalities.

The present study

In order to explore hallmark serial order STM phenomena in verbal and musical domains, we used serial order reconstruction tasks for auditorily presented sequences of verbal and musical material. Serial order reconstruction tasks allow maximizing the demands on serial order processing as item information is fully available during the reconstruction phase (see Majerus, Poncelet, Elsen, & van der Linden, 2006; Majerus, Poncelet, Van der Linden, & Weekes, 2008). Importantly, serial order reconstruction tasks further allow the assessment of musical and verbal STM abilities in a way that is not biased by output difficulties. Adult participants with no advanced musical expertise will be experts in outputting verbal stimuli, but they will be much less familiar with outputting musical stimuli. But even in musical experts, ease of musical output will depend on the type of the required musical response: professional singers will produce vocal musical responses with ease and high accuracy, but this will not necessarily be the case for professional pianists or percussionists. Therefore, the use of serial order reconstruction tasks—for which items are available at recall and only their order needs to be reconstructed—allows minimizing the impact of differential expertise levels for outputting verbal and musical responses on task performance. Also, in order to ensure comparable performance levels across the verbal and musical serial order reconstruction tasks, shorter lists were used for the musical than for the verbal task.

Finally, participants with no advanced musical expertise and experienced musicians were recruited for this experiment. This allowed us to determine whether any possible differences in serial order STM phenomena between verbal and musical STM tasks reflect fundamental serial order processing differences in the two domains, or whether they are the result of differences in expertise related to processing musical information. We furthermore expected that the experienced musicians group would show superior performance in the musical STM tasks as compared to the participants with no advanced musical expertise, in line with previous studies (Pechmann & Mohr, 1992; Schulze, Dowling, & Tillmann, 2012; Williamson, Baddeley, & Hitch, 2010).

Method

Participants. Eighty-eight participants with different levels of musical expertise took part in the present experiment on a voluntary basis; one participant had to be excluded due to data loss during data acquisition. The assignment of the participants to the experienced musician group or to the group

of participants with no advanced musical expertise was based on the participants' self-evaluation of their musical expertise. The participants were asked to rate their musical expertise using one of the following categories: 1 = non-musician; 2 = music-loving non-musician; 3 = amateur musician; 4 semiprofessional musician; 5 = professional musician (see Law & Zentner, 2012, for the use of similar catagories in determining musical achievement). Participants who chose one of the two first categories were assigned to the group of participants with no advanced musical expertise and the others to the musician group. The experienced musician group was composed of 41 participants (M_{age} = 23.0 years, SD = 4.0, 23 males; mean number of years of education, M = 13.9 years, SD = 1.7) and showed a high level of musical experience as expressed by the number of years of instrumental or singing practice (M = 11.5 years, SD = 6.0, range: 2-28). The group of participants with no advanced musical expertise was composed of 46 participants matched to the musician group in terms of age (M_{age} = 22.6 years, SD = 4.1, 11 males) and educational level (M = 14.1 years, SD = 2.0), and showing a very low level of musical experience (years of musical practice: M = 1.1 years, SD = 2.6, range: 0-12). The partially overlapping range of years of musical practice between the two groups is due to the fact that the experienced/nonexperienced musician categorization first depended on the participants' subjective estimation of their musical expertise. A small number of participants (n = 4) considered themselves as being experienced musicians while having a limited number of years of musical practice (3 years or less); there was also a slightly larger number of participants (n = 9) who considered that they were not expert musicians despite three years or more of musical practice. The two groups were also matched for nonverbal intellectual efficiency as assessed with the standard progressive Raven's matrices test (Raven, 1938). The matching of the two groups in terms of age, educational level, and nonverbal intelligence, was confirmed by Bayesian independent samples t tests assessing the evidence in favor of the null model (i.e., that there is no difference between the two groups). Evidence against a group effect was moderate for age (BF_{01} = 4.08) and educational level (BF_{01} = 3.96) and anecdotal for nonverbal intellectual efficiency (BF_{01} = 2.60), confirming overall the absence of a group effect on the matching variables. Finally, all the participants provided their written informed consent before starting the experiment, and the study had been approved by the Ethical

Board of the Faculty of Psychology, Speech and Language Therapy, and Education of the University of Liège.

Materials. The verbal and musical auditory sequences were constructed based on sets of nine digits and six tones, respectively. The spoken digits that we used ranged from 1 to 9 and were recorded by a French-speaking Belgian native male speaker (mean duration = 411 milliseconds, SD = 174). For the musical material, the tones consisted in the first six steps of a C major scale ranging from C_2 (65 Hertz) to A_2 (110 Hertz) and lasted for 300 milliseconds, with a fall and rise period of 10 milliseconds.

Design and procedure.

Verbal serial order reconstruction task. To assess verbal serial order STM, we used a serial order reconstruction task adapted from Majerus et al. (2008). Participants were administered digit lists of increasing length, ranging from six to nine items and presented in ascending order with four trials for each list length condition. Before the beginning of the task, participants were provided four trials of five-digit lists; these trials served as practice trials. The digits were presented at the pace of one item per second at a comfortable listening level via headphones connected to a portable workstation. For List Length N, the stimuli presented were the N first digits of the number sequence starting at 1. For example, six-digit lists were composed of digits from 1 to 6; seven-digit lists were composed of digits from 1 to 7, and so on. Immediately after the presentation of a trial, participants were given cards on which only the digits present in the trial were printed; the cards were aligned horizontally on the desk in numerical order. The participants had to reconstruct the original sequence by reordering the cards without any time constraint. The experimenter wrote down the participant's response and removed the cards. The participants were informed when sequence length increased via a message appearing on the screen.

Responses were scored for accuracy by determining the proportion of items output at their correct position and by averaging the score over the four trials for each sequence length. For serial position curve analyses, we determined for each serial position the proportion of items correctly output, separately for each sequence length. For the analysis of transposition gradients, we determined the proportion of serial position exchanges as a function of the distance of displacement. This analysis

was restricted to the longest sequence length (List Length 9), which produced the largest amount of serial position exchanges. Finally, for fill-in and infill errors we determined among all the trials the raw number of fill-in and infill errors. We considered an error as a fill-in error when an item was output one position too soon and was followed by the item that had occurred directly before this item in the target sequence (e.g., 1-3-2 instead of 3-1-2). An error was considered as an infill error when the same anticipation error was followed by the item following the anticipated item in the original sequence (e.g., 1-2-3 instead of 3-1-2).

Musical serial order reconstruction task. Participants were administered tone lists of increasing length (from three to six items) with four different trials per length condition. The administration of shorter sequences for the musical STM task, as compared to the verbal serial order reconstruction task, was motivated by the fact that STM capacities are overall lower for musical than for verbal stimuli, and this particularly true in nonmusician participants (Gorin, Kowialiewski, & Majerus, 2016; Schendel & Palmer, 2007; Schulze, Mueller, & Koelsch, 2011; Williamson, Baddeley, et al., 2010; Williamson, Mitchell, Hitch, & Baddeley, 2010). In order to ensure the familiarization with task requirements, participants were provided with three practice trials before starting the task. The first trial was performed by the experimenter showing the participants how to manipulate the reconstruction method (see below) for a two-tone list. Participants then performed two three-tone practice trials. The tone sequences were constructed and presented following the same procedure as for the verbal serial order reconstruction task; each trial was composed of the N first tones of the C_2 major scale, N being the list length level corresponding to a given trial. The tone sequences were presented at the rate of one tone every 360 milliseconds. Note that the presentation pace differed between the verbal and the musical tasks. By displaying musical stimuli at a faster pace, we aimed at ensuring that the sequence of the musical task were perceived as melodies and not as successions of isolated tones (for experiments using similar presentation rate, see, e.g., Dowling, 1991; Dowling, Bartlett, Halpern, & Andrews, 2008).

In order to reconstruct the tone sequences, participants used a virtual keyboard appearing on the screen immediately after sequence presentation (see Figure 1 for a graphical example). The virtual keyboard was composed of two rows of white circles shown on a black background; for each trial, the

number of circles displayed in each row was equal to the number of tones presented in the given trial. The circles in the first row represented the tones of the sequence, organized from the lowest to the highest tone. When participants clicked on a circle in the first row, the corresponding tone was played, and the circle became red, indicating that the tone was "active" (see Figure 1A). The participants could then click on a circle in the second row in order to indicate the serial position of the "active" tone in the memory sequence. When a position in the second row had been assigned to the "active" tone in the first layer, this position became green (assigned) and the tone in the first layer became gray (i.e., "used"; see Figure 2B). The participants could listen to the reconstructed tone sequence at any moment, and they had the possibility to change the order of the tones before validation, as for the verbal serial order reconstruction task. Also, as for the verbal task, no time constraint was imposed for sequence reconstruction responses. When a trial had been validated, the next trial started automatically, and participants were informed when sequence length was about to increase via a message appearing at the center of the screen, as was also the case for the verbal task. The scores and scoring methods were the same as for the verbal task, and the order of presentation of verbal and musical tasks, as well as the Raven's matrices test, was counterbalanced between participants.

Data and statistical analyses. Given recent criticisms relative to the use of frequentist statistical method when making statistical inferences (see, e.g., Dienes, 2011, 2016; Wagenmakers, 2007; Wagenmakers, Lee, Lodewyckx, & Iverson, 2008), all the statistical analysis conducted in the present study adopted a Bayesian approach. Bayesian statistical techniques have the advantage of relying on a model comparison rationale and adopting a model selection strategy. It also allows to compare H_1 and H_0 representing alternative and null models, respectively, and to select the model with the strongest evidence given the data.

All analyses report Bayes factors (BF), which can be considered as a relative measure of statistical evidence (Morey, 2015). The strength of evidence was interpreted as barely noticeable, moderate, strong, very strong, or decisive, when the BF was lesser than three, between three and 10, between 10 and 30, between 30 and 100, or higher than 100, respectively (see M. D. Lee & Wagenmakers, 2014). When reporting BFs, BF_{10} indicates the evidence for H_1 relative to H_0 and BF_{01}

indicates the reverse. Finally, all the analyses were conducted with Version 0.8.1.2 of the JASP

software package using default settings for Cauchy prior distribution (JASP Team, 2017).

Results[1]

    Sequence length.

    Verbal task. A first analysis assessed the overall effect of sequence length on response

accuracy. A 2×4 Bayesian mixed repeated-measures analysis of variance (ANOVA), with a two-

level between-participant factor (musicians vs. participants with no advanced musical expertise) and a

four-level within-participant list length factor (Length 6 to 9) showed that, after comparison to the

null model containing only the participant factor as nuisance variable, the model that received the

strongest evidence was the model containing the two main effects of group and of list length ($BF_{10} =$

7.26E+19). This model was followed by the model containing only the effect of list length ($BF_{10} =$

2.38E+19), which was 3.05 times less likely given the data. Overall, these results provide moderate

evidence that the data are better explained by a model containing both list length and participant

group effects, with decreasing performance as a function of increasing list length and superior

performance levels in the musician group (see Figure 2A). Finally, we checked that the recall

accuracy scores did not reflect random-level performance. We subtracted for each participant the

theoretical chance-level (.14) from his or her total accuracy score, and we compared this above-

chance score to a theoretical distribution centered on zero via a Bayesian one-sample *t* test. The

results provided decisive evidence in favor of above-chance-level performance ($BF_{10} = 2.86E+61$).

    Musical task. The same analysis was performed on response accuracy for the musical task

(see Figure 2B). A 2×4 Bayesian mixed repeated-measures ANOVA, with a two-level between-

participant factor (musicians vs. participants with no advanced musical expertise) and a four-level

within-participants list length factor (Length 3 to 6) showed that the model explaining the data best

was the model with the two main effects of group and of list length ($BF_{10} = 2.84E+22$). This model

was 6.55 times more likely than the model associated with the second-highest evidence, which

---

[1] To control for possible task order effects, the different analyses were also conducted by including a task order
variable (i.e., verbal task first / musical task first). These additional analyses revealed no effect of, or interaction
with, the task order variable across all the analyses.

included all main effects plus their interaction (BF_{10} = 4.33E+21). These results therefore provide

clear evidence in favor of the presence of both sequence length and musical expertise effects.

We furthermore checked that the lower performance in the participant group with no

advanced musical expertise was not due to floor-level performance. We calculated the theoretical

chance-level for performance (.24), and we subtracted it from each participant's total accuracy score.

When comparing these chance-level corrected accuracy scores to a theoretical distribution centered on

zero via a Bayesian one-sample t test, we observed decisive evidence for above-chance-level

performance in both the participant group with no advanced musical expertise (BF_{10} = 1.14E+6) and

the musician group (BF_{10} = 3.65E+15).

Serial position curve.

Verbal task. We conducted, separately for each list length, Bayesian mixed repeated-

measures ANOVAs with a two-level between-participant factor (musicians vs. participants with no

advanced musical expertise) and a within-participant serial position factor (from 1 to n levels, n

depending on list length) on verbal serial order reconstruction accuracy scores. The results obtained

for the different sequence lengths were very similar. For six-digit and eight-digit sequences (see Table

1), two models received similar levels of large evidence compared to the null model: the model with

only the effect of serial position, and the same model including also the participant group effect. For

seven-digit and nine-digit lists (see Table 1), the two models receiving similar large levels of evidence

compared to the null model were the model including both serial position and group effects, and the

model including only the serial position effect. Rouder, Engelhardt, McCabe, and Morey (2016)

proposed that in the context of a model comparison approach, the simpler model should always be

preferred when two models lead to comparable levels of evidence. In the present case, the simpler

model including only the serial position factor should thus be preferred for all sequence length

conditions. This was confirmed by an analysis of specific effects, which averages the evidence for the

effect of interest across all the models containing the effect. The serial position was associated with

decisive evidence ($BF_{Inclusion}$ = ∞, for all sequence lengths), while evidence for an effect of musical

expertise was very low (Length 6: $BF_{Inclusion}$ = 0.47; Length 7: $BF_{Inclusion}$ = 1.19; Length 8: $BF_{Inclusion}$ =

0.50; Length 9: $BF_{Inclusion}$ = 0.63). Note that this analysis, assessing performance separately for each

list length, did not confirm the group effect observed when performance was assessed over the entire

task; the group effect for the verbal serial order reconstruction task hence needs to be considered with

caution.

Next, we assessed more directly the presence of primacy and recency effects by collapsing

data across groups. Bayesian paired-samples t tests compared recall accuracy between Positions 1 and

2 (primacy effect), as well as between the last and penultimate positions (recency effect). We obtained

moderate to decisive evidence for the presence of a primacy effect in each list length (Length 6: BF_{10}

= 80.07; Length 7: BF_{10} = 3.93; Length 8: BF_{10} = 5.72E+3; Length 9: BF_{10} = 2.24E+3). We also

obtained decisive evidence for the presence of a recency effect in each list length (Length 6: BF_{10} =

2.32E+8; Length 7: BF_{10} = 1.58E+7; Length 8: BF_{10} = 6.37E+12; Length 9: BF_{10} = 1.88E+7). Figure

3 shows clear recency and primacy effects for all list lengths in the verbal serial order reconstruction

task.

Musical task. The same type of analyses considering response accuracy as a function of serial

position was conducted for the musical serial order reconstruction task. As shown in Table 2, for

three-tone and four-tone sequences, the model associated with largest evidence, relative to the null

model, was the model containing only the effect of participant group. As shown in Figure 4A and 4B,

the response curves were flat, probably due to the limited number of items that were presented for

these list length (see also Smyth et al., 2005, for similar results for a three-item visuospatial serial

order reconstruction task). For five-tone and six-tone sequences, recency effects were observed, as

shown in Figure 4C and 4D. For both list lengths, the model receiving the strongest evidence was the

model including both serial position and group effects (see Table 2). The presence of a serial position

effect was supported by an analysis of specific effects providing moderate and decisive evidence for

the effect of serial position in five-tone ($BF_{Inclusion}$ = 3.17) and 6-six-tone sequences ($BF_{Inclusion}$ =

220.56), respectively. Decisive evidence was also obtained for the presence of an effect of musical

expertise for five-tone ($BF_{Inclusion}$ = 3.60E+3), and six-tone sequences ($BF_{Inclusion}$ = 2.03E+3).

Next, we directly assessed primacy and recency effects in List Lengths 5 and 6; as for the

verbal tasks, data were collapsed across groups, and this was further motivated by the fact that there

was no evidence for a group-by-serial position interaction. The presence of a recency effect in List

Length 5 was supported by very strong evidence (comparison of Positions 4 and 5: $BF_{10} = 36.02$).

However, there was no evidence for a primacy effect (comparison of positions 1 and 2, see Figure

4C), with evidence even slightly in favor of the null effect ($BF_{01} = 1.60$). For List Length 6, we

obtained strong evidence for a recency effect (comparison of Positions 5 and 6: $BF_{10} = 18.80$) but no

evidence for a primacy effect (comparison of Positions 1 and 2); evidence was slightly in favor of the

null effect for the primacy effect ($BF_{01} = 1.15$). At the same time, for List Length 6, when inspecting

the form of the serial response curve presented in Figure 4D, a primacy effect appears to be visible,

but only when comparing Position 1 to later serial positions, such as Serial Position 3, indicating a

more shallow primacy portion as compared to the verbal task. A direct comparison between Positions

1 and 3 provided indeed decisive evidence for a recall superiority of Position 1 versus Position 3

($BF_{10} = 614.67$), and moderate evidence for a recall superiority of Position 2 versus Position 3 ($BF_{10} =$

3.13). In sum, recency and primacy effects characterize both verbal and musical serial order

reconstruction tasks, but at the same time, these effects, and particularly the primacy effect, are less

pronounced for the musical task, even when comparing lists of identical length for both tasks (List

Length 6).

      Transposition gradients.

      Verbal task. The next analyses assessed transposition gradients by analyzing serial position

errors as a function of displacement distance and by limiting this analysis to the list length for which

the largest number of serial position errors had been observed (List Length 9 for the verbal task, List

Length 6 for the musical task). For the verbal serial order reconstruction task, we conducted a

Bayesian mixed repeated-measures ANOVA on serial position errors with a two-level between-

participant factor (musicians vs. participants with no advanced musical expertise) and a seven-level

within-participant factor of absolute displacement distance (Distances 1 to 7[2]). The results[3] showed

that, compared to the null model, the model associated with the strongest evidence was the model

including only the effect of displacement distance ($BF_{10} = 5.23E+116$; see Figure 5A), followed by

---

[2] Since no transposition errors of Distance 8 were observed in any participant, only Distances 1 to 7 were
included in the analysis.
[3] Note that the data of three musician participants and one participant with no advanced musical expertise were
excluded from analysis, as these participants did not show any serial position errors for list length 9.

the model with the effects of displacement distance and group (BF_{10} = 5.59E+115), the latter being

9.37 times less likely than the former. These results replicate previous findings in the verbal STM

domain, by showing that the proportion of serial order errors decreases as a function of displacement

distance, and this similarly in musician and nonmusician participants.

   Musical task. The same analysis of transposition gradients was conducted on the serial order

errors for the musical serial order reconstruction task. We conducted a Bayesian mixed repeated-

measures ANOVA, with a two-level between-participant factor and a five-level within-participant

factor of absolute displacement distance (Distances 1 to 5). The results[4] showed that compared to the

null model, the full model containing all the main effects and their interaction received the strongest

evidence (BF_{10} = 5.83E+61), followed by the model with only the main effect of displacement

distance (BF_{01} = 8.95E+60), which was 6.51 less likely. In order to explore the Group $\times$ Displacement

Distance interaction, pairwise comparisons between the different displacement distances were

conducted within each group, using Bayesian paired-samples *t* tests. For participants with no

advanced musical expertise, pairwise comparisons revealed a progressive decrease of serial order

errors as a function of displacement distance increase (see Figure 5B; Distance 1 > Distance 2: BF_{10} =

6.31; Distance 2 > Distance 3: BF_{10} = 2196.33; Distance 3 > Distance 4: BF_{10} = 10.64; Distance 4 >

Distance 5: BF_{10} = 562.89). However, a different pattern of transpositions was observed for the

musician group. There was no evidence for a decrease of serial order errors for Distance 1 versus

Distance 2 displacements (BF_{10} = 0.18); the results actually moderately favored the absence of

difference between the two distances (BF_{01} = 5.59). The proportion of serial order errors, however,

decreased between Distance 2 and Distance 3 (BF_{10} = 2.36E+5), with no further difference for all

successive distances (Distance 3 vs. Distance 4, BF_{10} = 0.37; Distance 4 vs. Distance 5, BF_{10} = 0.65;

evidence for an absence of an effect: Distance 3 vs. Distance 4, BF_{01} = 2.67; Distance 4 vs. Distance

5, BF_{01} = 1.54).

---

[4] Note that the data of one musician participant were excluded from analysis, as this participant did not show
any transposition errors during the whole task.

It appears that while participants with no advanced musical expertise showed a linear decrement of transposition errors as a function of the increase of displacement distance in both verbal and musical tasks, this was only the case in musician participants for the verbal task.

Fill-in:infill errors ratio.

Verbal task. In a final analysis, we examined the occurrence of fill-in and infill errors in the verbal and musical serial order reconstruction tasks by collapsing the different trial lengths. A Bayesian mixed repeated-measures ANOVA, with a two-level between-participant factor (musicians vs. participants with no advanced musical expertise) and a 2-level within-participant type of error factor (fill-in vs. infill) showed[5] that, compared to the null model, the model with the effect of error type received the strongest evidence ($BF_{10} = 1.46E+17$) and was 2.78 more likely than the model containing both the group and type of error effects ($BF_{10} = 5.27E+16$). Complementary analyses of specific effects confirmed the presence of an error type effect ($BF_{Inclusion} = 4.62E+14$), while the effect of group received very little evidence ($BF_{Inclusion} = 0.29$). The effect of error type was characterized by more fill-in than infill errors (ratio = 2.89; see Figure 6A).

Musical task. The same analysis[6] was conducted for the musical task and revealed that compared to the null model, the most likely model was the full model containing both main effects and the interaction ($BF_{10} = 4.69E+5$). The second best model was the model containing the main effects of type of error and of group ($BF_{10} = 352.54$) and was 1329.68 times less likely.

An exploration of the interaction via pairwise Bayesian *t* tests provided strong evidence for a reduced proportion of fill-in errors in musicians as compared to participants with no advanced musical expertise ($BF_{10} = 646.96$, see Figure 6B). An absence of group effect was observed for the amount of infill errors ($BF_{10} = 0.28$; $BF_{01} = 3.60$). Furthermore, participants with no advanced musical expertise produced more fill-in than infill errors with a ratio of 2.32 ($BF_{10} = 1.15E+4$; see Figure 6B), but this was not the case for musicians (fill-in:infill ratio = 0.94; $BF_{10} = 0.18$; $BF_{01} = 5.46$).

---

[5] Note that the data from one participant were discarded from analysis, as this participant did not make any error during the whole task; this participant was also among those excluded in the analyses of transposition gradients of the verbal task.

[6] Note that the data from one participant were excluded from analysis, as this participant did not show any error during the whole task; this participant was also among those excluded in the analyses of transposition gradients of the musical task.

Discussion

This study aimed at investigating hallmark serial order effects across verbal and musical domains of STM. We observed similar serial order phenomena in musical and verbal short-term serial order reconstruction tasks, as highlighted by similar transposition gradients, fill-in:infill errors ratios, and sequence length effects. Recency and primacy effects also characterized recall performance in both tasks, but for musical STM tasks these effects only appeared for the longest list lengths and were less marked. Finally, in addition to outperforming participants with no advanced musical expertise on the musical STM task, musicians showed a specific pattern of serial recall errors in the musical task. Indeed, relative to the group with no advanced musical practice, musicians showed an increase of Distance-2 transposition errors and a similar amount of fill-in and infill errors.

We observed similar sequence length effects in the two STM domains, with performance decreasing as a function of increasing sequence length, in line with previous studies having shown robust list length effects across different STM domains (e.g., Anderson, Bothell, Lebiere, & Matessa, 1998; Avons, 1998; Jones et al., 1995; Maybery, Parmentier, & Jones, 2002; Smyth et al., 2005; Williamson, Baddeley, et al., 2010). As regards serial position effects, we observed in the verbal task a classical U-shaped serial curve accompanied by both primacy and recency effects, and this for the four different list lengths that were explored, in accordance with many other studies from the verbal, musical and visuospatial domains (Avons, 1998; Cowan et al., 2002; Greene & Samuel, 1986; Guérard & Tremblay, 2008; Mondor & Morin, 2004; Oberauer, 2003; Smyth et al., 2005; Tremblay, Parmentier, Guérard, Nicholls, & Jones, 2006). For the musical task, we also observed primacy and recency effects, but only for the longest list lengths, in both musicians and participants with no advanced musical expertise. These effects, when present, were less marked as compared to the verbal task.

For the smallest musical STM lists—that is, three-tone and four-tone sequences—serial response curves were flat. This result is, however, non-informative given that it is very likely to be caused by the small number of positions used (see also Smyth et al., 2005, for similar results). The absence in the musical task of a primacy effect for List Length 5, and the relatively weaker primacy effect for List Length 6 as well as the weaker recency effects for List Lengths 5 and 6, deserves some

more consideration. A possible explanation for these relatively flat serial position curves is that

participants may have relied on configurational melodic processing instead of strict serial processing

of the memory items (e.g., contour information reflects the overall pattern of up and down tone

transitions in the sequence, and can be represented in a gestalt-like manner; see Dowling, 1991; Gorin

et al., 2016). In the musical literature, it is known that contour information plays an important role in

short-term recognition of melodic excerpts (Dowling, 1978; Dowling & Tillmann, 2014). For

example, consider the tone sequence C_2–E_2–D_2–F_2 which is characterized by a melodic contour of two

rise–fall alternations (resembling, in spatial terms, to a landscape with two horizontally aligned

mountains in the background). If for the maintenance of this sequence a melodic contour

representation is used, then errors are more likely to involve the exchange of C_2 and D_2, or E_2 and F_2,

leading to a preservation of the rise–fall pattern of the melodic contour. This view is clearly supported

by the presence in four-tone sequences of more transpositions for Distance 2 than for Distance 1 and

Distance 3 (see Figure 7A). At the same time, in five-tone sequences, the proportion of transpositions

decreased linearly as a function of distance of displacement (see Figure 7B), which suggests that this

type of configurational representation was not used anymore for longer musical sequences.[7]

     This potential shift from a gestalt-like contour representation to a more serial representation is

in line with other studies in the musical STM domain. Edworthy (1985) showed that participants are

better at recognizing contour information in short sequences and better at discriminating more fine-

grained pitch-interval information in longer sequences. In order to check for the intervention of these

possible strategies, we reanalyzed our data by scoring the responses based on contour recall, by

considering two adjacent tones as correct if their interval direction (up or down) conformed to the

contour of the sequence. This score was then corrected for chance-level performance and compared to

the strict serial order score (also after chance-level correction as reported in the Results section), by

collapsing three-tone and four-tone lists, on the one hand, and five-tone and six-tone lists, on the other

hand, in order to maximize the reliability of these analyses. A mixed Bayesian repeated-measures

---

[7] Note that, since some participants made no errors for List Lengths 4 and 5, panels A and B in Figure 7 displays transposition gradients of 44 "non-musicians" and 28 musician participants, and 46 "non-musicians" and 36 musician participants, respectively.

ANOVA (2 groups $\times$ 2 scores $\times$ 2 list length) showed strong evidence for a Score $\times$ Length interaction ($BF_{Inclusion} = 13.16$), suggesting that contour and serial position scoring methods led to opposing results as a function of sequence length. Furthermore, there was some evidence for a three-way interaction ($BF_{Inclusion} = 2.29$), indicating that the Score $\times$ Length interaction was further modulated by group. As shown in Figure 8A, the musician group showed better recall performance when the serial position scoring method was applied, and this was for all sequence lengths. For participants with no advanced musical expertise, the contour recall score captured performance to the same extent or even better than the serial recall score for the short list lengths, while the serial recall score led to higher performance for the longer list lengths, in line with a gradual shift from contour-based to serial position-based representations.

Another explanation for the presence of weak primacy and recency effects in five-tone and six-tone lists could be related to the use of faster presentation rates for the musical stimuli as compared to the verbal stimuli. A fast presentation rate had been chosen for the musical stimuli as our initial aim was to ensure that they were perceived as melodies rather than as series of isolated tones and that they were not verbally recoded. At the same time, it has been shown in the verbal domain that presentation rates can influence the shape of the serial position curve (see, e.g., Posner, 1964; Tan & Ward, 2008). Fast presentation rates of verbal stimuli have been shown to lead to reduced primacy effects (Bhatarah, Ward, Smith, & Hayes, 2009). Conversely, more pronounced primacy and recency effects have been shown in musical STM tasks requiring recall of tone sequences presented at slower rates (i.e., one tone per second; Greene & Samuel, 1986). At the same time, another factor known to influence the shape of the serial position curve is list length. Indeed, previous studies showed that when using a reconstruction method for short lists of verbal and visual memoranda, recency and primacy effects can be reduced, leading to a flattened shape of the serial response curve (Smyth et al., 2005; Ward, Tan, & Grenfell-Essam, 2010). Also, studies on free recall of spatial locations argued that the tendency to serially recall items according to their initial list position could be language-specific (see, e.g., Gmeindl, Walsh, & Courtney, 2011). However, other recent studies showed that serial recall strategies can be observed in free recall tasks for nonverbal memoranda such as spatial or facial locations (Cortis, Dent, Kennett, & Ward, 2015). Future studies using similar presentation rates

and list lengths when comparing performance in musical and verbal STM tasks are required to determine whether the differences in serial position curves observed in this study for verbal and musical STM tasks are due to differences in task parameters or to the existence of specific serial order coding processes in the musical domain (such as the use of contour-based configurational processes).

We also observed that, as compared to participants with no advanced musical expertise, musicians produced specific patterns of errors in the musical serial order reconstruction STM task, such as specific transposition gradients and fill-in:infill ratios. This suggests that musicians may use additional domain-specific strategies for processing musical information (see also Williamson, Baddeley, et al., 2010). One critical variable for efficiently processing musical stimuli is meter. Meter refers to the generation of a metrical hierarchy taking the form of a recurrent alternation between strong and weak beats occurring at different periodicities (Lerdahl & Jackendoff, 1983). Musicians possess elaborate metrical mental representations they will use to structure novel musical sequences. The use of metrical representations will allow the chunking of musical elements, with an effect similar to temporal grouping. This is likely to have an effect on STM performance and error types. Mathias et al. (2015) observed that pianists produced more transpositions between items having the same metrical role, particularly in musical contexts with a stronger metrical structure. In the present case, the use of an elementary, binary metrical structure (e.g., strong, weak alternation) could have influenced the pattern of errors by enhancing the amount of serial position exchanges between positions sharing the same metrical role (i.e., every $n+2$ position). This could explain why musicians showed an increase of Distance 2 transposition errors and an altered fill-in:infill ratio, relative to participants with no advanced musical expertise.

At the same time, these specificities characterized mainly performance in musicians. For participants with no advanced musical expertise, serial order effects were very similar across verbal and musical modalities (except for the reduced primacy and recency effects in the musical task). Indeed, analysis of transposition gradients showed that transposition errors were governed by a locality principle (Henson, 1996) in both modalities; the same transposition gradients have also been observed for verbal and visuospatial STM tasks (Avons, 1998; Farrell & Lewandowsky, 2004; Hartley et al., 2016; Henson, 1998; Hurlstone & Hitch, 2015; Johnson, Shaw, & Miles, 2016;

Parmentier, Andrés, et al., 2006; Parmentier, King, et al., 2006; Smyth et al., 2005). Also, fill-in errors

were overall 2 times more frequent than infill errors in verbal and musical modalities (Farrell et al.,

2013; Henson, 1996; Surprenant et al., 2005), mirroring again results also observed in the visuospatial

(Guérard & Tremblay, 2008) and tactile domains (Johnson et al., 2016).

What do these serial order effects tell us about the nature of ordering mechanisms in musical

STM? The presence in participants with no advanced musical expertise of a very similar profile of

serial order behaviors in musical and verbal STM strengthens the view that musical and verbal STM

systems possibly rely on common sequential processes (Gorin et al., 2016; Williamson, Baddeley, et

al., 2010). These processes have been proposed to rely on domain-general serial order marking

processes, such as context signals (see Hurlstone et al., 2014; Lewandowsky & Farrell, 2008). The

presence in the musical task of sequence length effect, transposition gradients, and serial position

effects fit with positional marking (e.g., Brown et al., 2000; Burgess & Hitch, 1999; Hartley et al.,

2016) and primacy gradient (e.g., Farrell & Lewandowsky, 2002; Page & Norris, 1998) mechanisms

as proposed by verbal STM models. Moreover, the presence of more fill-in than infill errors (see

Henson, 1996; Surprenant et al., 2005) provides evidence for the intervention of a primacy gradient

principle to represent serial order information in musical STM while arguing against a chaining

account (e.g., Kieras et al., 1999; Lewandowsky & Murdock, 1989). However, even though we

obtained in participants with no advanced musical expertise evidence in favor of similar ordering

mechanisms in verbal and musical STM domains, we cannot rule out the possibility that distinct,

modality-specific STM systems process serial order information using similar representational

mechanisms (see Logie et al., 2016). At the same time, the results observed in this study can be

accommodated by an approach considering STM as an emerging process resulting from the

coactivation of domain-specific systems for storage of verbal and musical item information, and

domain-general mechanisms involved in serial order coding and attentional control (see, e.g.,

Majerus, 2013; Majerus et al., 2016).

The effects observed here stress the need for musical STM models to take into consideration

the problem of serial order. There is currently no model of musical STM/working memory addressing

directly this aspect of musical sequence processing. In his model, Berz (1995) suggested to add to the

Baddeley and Hitch (1974) multicomponent model of working memory a module specialized in the processing of musical stimuli. However, as in the original working memory model, the model of Berz (1995) lacks details concerning the representation and the processing of serial order information. More recently, Ockelford (2007) suggested the existence of a central executive component specialized in the processing of musical material that is connected to short-term and long-term musical memory stores. Moreover, Ockelford (2007) proposed that when experiencing music, the musical executive component is responsible for tagging musical elements as well as their relationships. This tagging is proposed to rely on item-by-item associative chaining mechanisms, following the proposal of Kieras et al. (1999). However, as we have seen, chaining mechanisms are not compatible with the observation of more fill-in than infill errors. Future models of musical STM will need to incorporate realistic mechanisms for explaining serial ordering phenomena in the musical domain, and, as shown by the present data, some of these mechanisms may prove to be the same as those used in the verbal STM domain, such as context signals based on primacy gradients and/or temporal/episodic signals (Brown et al., 2000; Burgess & Hitch, 1999; Hartley et al., 2016; Page & Norris, 1998). At the same time, musical STM models also need to account for the more specific processes that characterize musical serial order processing such as the possible influence of metrical knowledge and structure, as well the use of contour-based representations. Hence, in addition to domain-general position marking and primacy gradient mechanisms, models of musical STM also need to incorporate metrical and configurational levels of processing and their interaction with serial position coding.

The present study has also methodological implications. Our results indicate that the reconstruction of the serial order of tone sequences via a virtual keyboard is feasible by musically trained participants but also by participants with no advanced musical expertise. This new method thus offers the possibility to study STM for serial order in both verbal and musical modalities, more particularly in nonmusicians. Nonetheless, it is important to acknowledge that the musical reconstruction task that we used does not perfectly match the verbal reconstruction task used in this study. In the verbal task, the memoranda were printed on cards and were directly available during serial order reconstruction, and, hence, participants had permanent access to item identity. In the musical task, participants needed to reactivate each tone by a clicking response in order to hear its

identity. These methodological discrepancies could partly explain the differences in the shapes of serial position curves we observed between the verbal and the musical tasks, as discussed above.

One manner to address this issue could be to use direct recall procedures for the tasks in both the verbal and musical domains. At the same time, previous studies have shown that imprecise singing is widespread in the general population, and even people considered as accurate singers do not imitate perfectly pitch target (see Pfordresher & Brown, 2007; Pfordresher, Brown, Meier, Belyk, & Liotti, 2010). Also, the accuracy of sung productions is assessed through relative measures of deviation, determining the extent to which a produced pitch or interval deviates from its target (see Berkowska & Dalla Bella, 2013); these measures are difficult to implement in tone serial recall tasks. The use of a threshold above which the deviation of a produced pitch is considered as incorrect could represent a possibility to circumvent this methodological challenge. However, another problem raised by this method concerns the difficulty of differentiating between item-based and order-based errors as deviating tones could be both item and order errors, depending on the target item to which they are compared (i.e., a given tone could deviate by a semitone relative to the stimulus presented in the same position, leading to an item error, or by a semitone relative to a stimulus in another serial position, which would be considered as a serial order error). Another possibility could be to rely on identical probe recognition procedures for assessing verbal and musical STM. At the same time, recognition paradigms are not suited for studying serial order errors given that participants respond simply with yes–no responses. The serial order errors a participant could have made could be inferred by the presentation of nonmatching probe stimuli containing these serial order errors, but this approach would be very indirect and of suboptimal sensitivity and efficiency. Recognition STM paradigms are indeed better suited to study recognition accuracy (Henson, Hartley, Burgess, Hitch, & Flude, 2003; Jefferies, Frankish, & Lambon Ralph, 2006). The use of a reconstruction method provides a more direct approach for studying serial recall accuracy as well as ordering errors in the musical domain. Moreover, the present study indicated that this type of task is a valid tool for measuring musical STM for serial order in both musician and nonmusician participants, allowing for generalization of the results to the general population.

Finally, a further finding of this study is the musicians' overall higher musical STM performance and, to some extent, their higher verbal STM performance. Concerning musical STM, our results are in line with previous studies reporting that musicians have better abilities to keep in memory pitch information (Berti, Münzer, Schröger, & Pechmann, 2006; Pechmann & Mohr, 1992) and to recall and recognize the serial order of tone sequences (Schulze et al., 2012; Williamson, Baddeley, et al., 2010). Other studies observed that this advantage also extends to the verbal domain (Chan, Ho, & Cheung, 1998; Franklin et al., 2008; Hansen, Wallentin, & Vuust, 2013; Y. Lee, Lu, & Ko, 2007; e.g., Parbery-Clark, Skoe, Lam, & Kraus, 2009). The present study yielded some evidence for such an advantage, even though the advantage for STM performance in the verbal modality was less reliable than the advantage for STM performance in the musical modality. Moreover, it is important to note that we did not directly control for differences in attentional capacities between musicians and participants with no advanced musical expertise, although both groups were matched for nonverbal intellectual efficiency. A number of studies have shown that musicians not only have better STM abilities but also more developed auditory attentional capacities (see Besson, Chobert, & Marie, 2011; Strait, Kraus, Parbery-Clark, & Ashley, 2010). Attentional focalization is an important process involved in many different STM tasks, if not all (Barrouillet, Bernardin, & Camos, 2004; Cowan, 1995); and higher auditory attentional focalization abilities in musician participants could at least partially explain the small advantage observed for verbal STM performance.

Also, it could be argued that the distinction we made between the two groups—that is, musically trained participants versus participants with no advanced musical expertise—is somewhat simplistic. It has been shown that individuals considered as musicians can show surprisingly low levels of performance in various tests assessing musical abilities, while other individuals considered as nonmusicians can show high performance levels on the same tests (see Law & Zentner, 2012). Likewise, even though studies on musical aptitude have shown that musical test batteries can estimate a general form of musical aptitude in both musician and nonmusician participants, they also have highlighted the existence of varying patterns of musical sophistication, with different musical aspects (such as pitch-related vs. rhythm-related skills) being developed to different extents (see, e.g., Kunert, Willems, & Hagoort, 2016; Law & Zentner, 2012; Müllensiefen, Gingras, Musil, & Stewart, 2014).

Therefore, it is important that future studies interested in the link between STM capacities and level of musical sophistication take into consideration not only the musician versus nonmusician distinction but also the extent to which participants are proficient in various musical skills, such as, for example, assessment with the short version of the Profile of Music Perception Skills battery (Law & Zentner, 2012). For instance, a close link between serial order STM and rhythm processing has been observed (Gorin et al., 2016; Plancher, Lévêque, Fanuel, Piquandet, & Tillmann, 2017; Saito, 2001). Given that temporal aspects such as rhythm play a key role in some models of verbal STM for serial order (see Brown et al., 2000; Burgess & Hitch, 1999; Hartley et al., 2016), it would be theoretically useful to investigate, using an individual differences approach, the relation between rhythmic aptitudes and STM for serial order.

To sum up, this study observed similar serial order phenomena for verbal and musical STM tasks in participants with no advanced musical expertise suggesting, that domain-general mechanisms support the representation of serial order information in STM. At the same time, this study also highlights domain-specific serial processes, possibly involving configurational and metrical processes, most clearly in participants with advanced musical experience. Future models of STM need to consider both domain-general and domain-specific serial order coding mechanisms.

Table 1

Statistical results for Bayesian mixed repeated measures ANOVAs on reconstruction accuracies for

six-digit to nine-digit sequences

Length	Models	P(M)	P(M|data)	BF_M	BF_{10}	% error
6-digit	Null model (incl. subject)	0.200	1.976E-25	7.903E-25	1.000	
	Position	0.200	0.589	5.727	2.980E+24	0.489
	Group	0.200	1.161E-25	4.642E-25	0.587	1.044
	Position + Group	0.200	0.391	2.565	1.977E+24	1.119
	Position + Group + Position × Group	0.200	0.021	0.084	1.040E+23	1.028
7-digit	Null model (incl. subject)	0.200	8.133E-21	3.253E-20	1.000	
	Position	0.200	0.358	2.235	4.407E+19	0.461
	Group	0.200	1.287E-20	5.114E-20	1.572	2.836
	Position + Group	0.200	0.618	6.475	7.600E+19	3.283
	Position + Group + Position × Group	0.200	0.023	0.096	2.881E+18	2.358
8-digit	Null model (incl. subject)	0.200	7.098E-41	2.839E-40	1.000	
	Position	0.200	0.570	5.310	8.036E+39	0.447
	Group	0.200	4.752E-41	1.901E-40	0.669	1.757
	Position + Group	0.200	0.426	2.973	6.006E+39	0.741
	Position + Group + Position × Group	0.200	0.003	0.013	4.660E+37	0.869
9-digit	Null model (incl. subject)	0.200	2.152E-35	8.607E-35	1.000	
	Position	0.200	0.513	4.209	2.383E+34	0.304
	Group	0.200	1.342E-35	5.367E-35	0.624	1.120
	Position + Group	0.200	0.381	2.459	1.769E+34	1.648
	Position + Group + Position × Group	0.200	0.107	0.477	4.952E+33	0.889

Note. All models include subject variable.

Table 2

Statistical results for Bayesian mixed repeated measures ANOVAs on reconstruction accuracies for three-tone to six-tone sequences

| Length	Models	P(M)	P(M|data)	BF_M	BF_{10}	% error
3-tone	Null model (incl. subject)	0.200	1.263E-4	5.053E-4	1.000	
	Position	0.200	9.158E-6	3.663E-5	0.072	1.570
	Group	0.200	0.917	44.001	7256.765	1.096
	Position + Group	0.200	0.066	0.281	519370	1.309
	Position + Group + Position × Group	0.200	0.018	0.072	139.255	1.899
4-tone	Null model (incl. subject)	0.200	1.676E-6	6.706E-6	1.000	
	Position	0.200	2.733E-8	1.093E-7	0.016	0.744
	Group	0.200	0.983	227.941	5.86E+5	1.773
	Position + Group	0.200	0.016	0.067	9759.002	1.508
	Position + Group + Position × Group	0.200	8.831E-4	0.004	526.763	6.086
5-tone	Null model (incl. subject)	0.200	4.450E-5	1.780E-4	1.000	
	Position	0.200	1.406E-4	5.624E-4	3.16	0.923
	Group	0.200	0.174	0.84	3899.408	0.726
	Position + Group	0.200	0.564	5.169	1.27E+4	1.803
	Position + Group + Position × Group	0.200	0.263	1.424	5901.347	1.476
6-tone	Null model (incl. subject)	0.200	1.008E-6	4.033E-6	1.000	
	Position	0.200	3.264E-4	0.001	323.803	0.475
	Group	0.200	0.003	0.012	2988.123	4.383
	Position + Group	0.200	0.966	113.272	9.58E+5	1.813
	Position + Group + Position × Group	0.200	0.031	0.127	3.05E+4	1.701

Note. All models include subject variable.
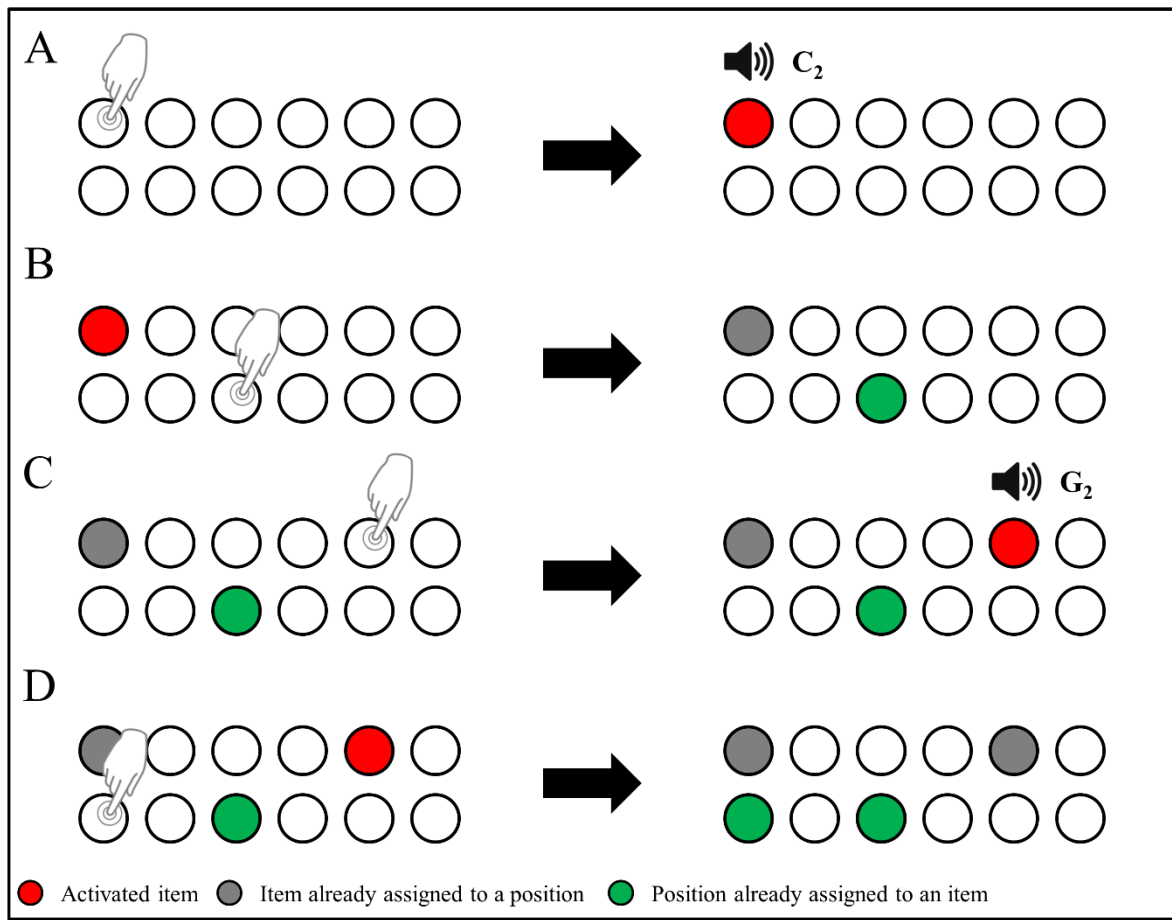
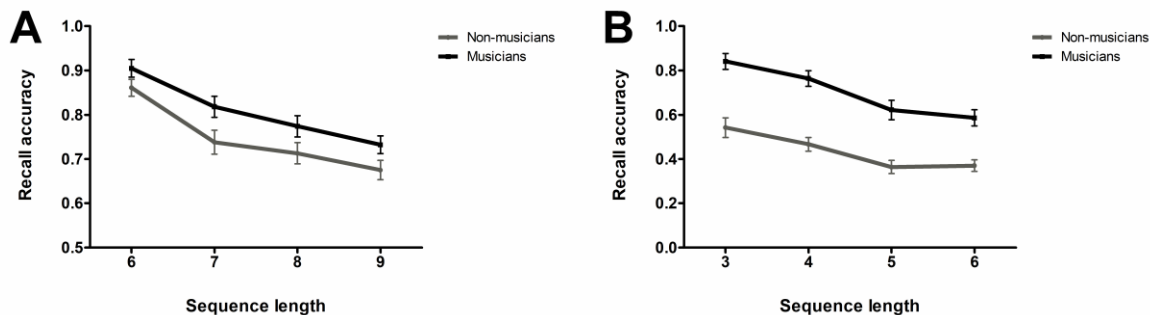*Figure 1*. Graphical representation of the musical serial order reconstruction task.

Figure 2. (A) Means and standard errors for response accuracy in the verbal order reconstruction task, as a function of sequence length (from 6 to 9) and participant group (musicians vs. participants with no advanced m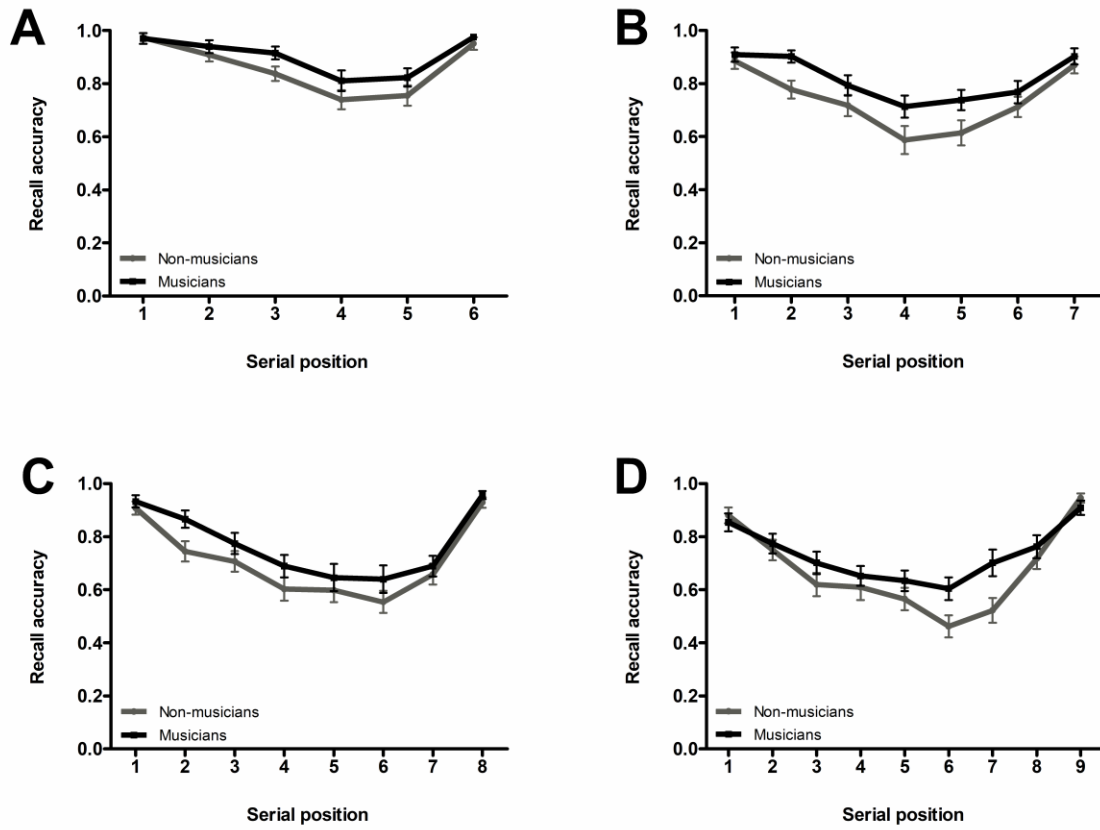usical expertise, referred to as "non-musicians"). (B) Means and standard errors for response accuracy in the musical order reconstruction task, as a function of sequence length (from 3 to 6) and participant group (musicians vs. participants with no advanced musical expertise referred to as "non-musicians").

Figure3. Means and standard errors for response accuracy in the verbal order reconstruction task, as a function of serial position, participant group (musicians vs. participants with no advanced musical expertise, referred to as "non-musicians"), and sequence length: (A) Sequence Length 6, (B) Sequence Length 7, (C) Sequence Length 8, and (D) Sequence Length 9.
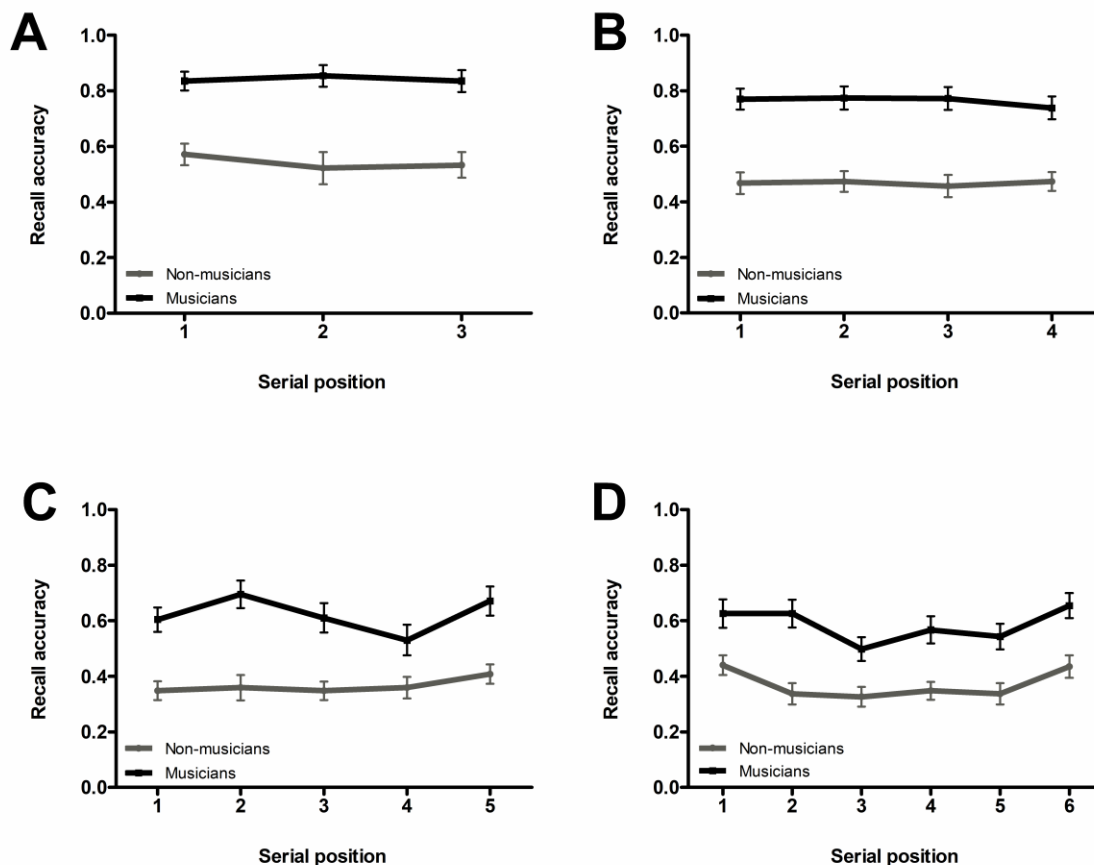
Figure 4. Means and standard errors for response accuracy in the musical order reconstruction task, as a function of serial position, participant group (musicians vs. participants with no advanced musical expertise, referred to as "non-musicians"), and sequence length: (A) Sequence Length 3, (B) Sequence Length 4, (C) Sequence Length 5, and (D) Sequence Length 6.
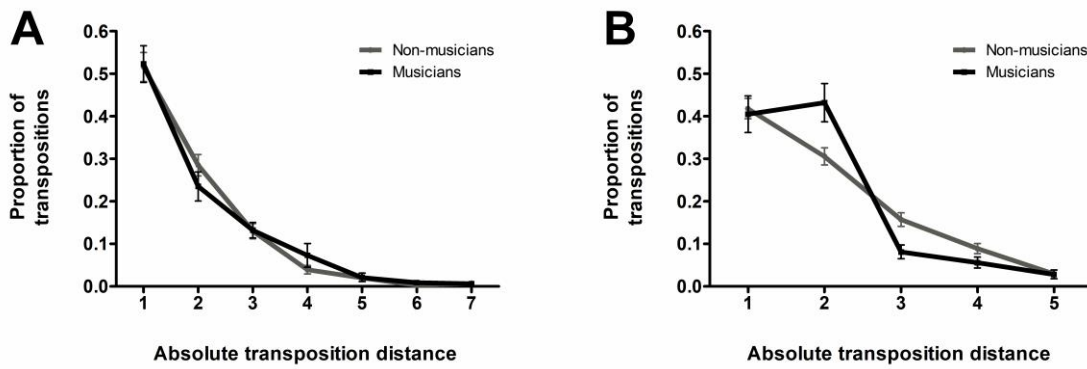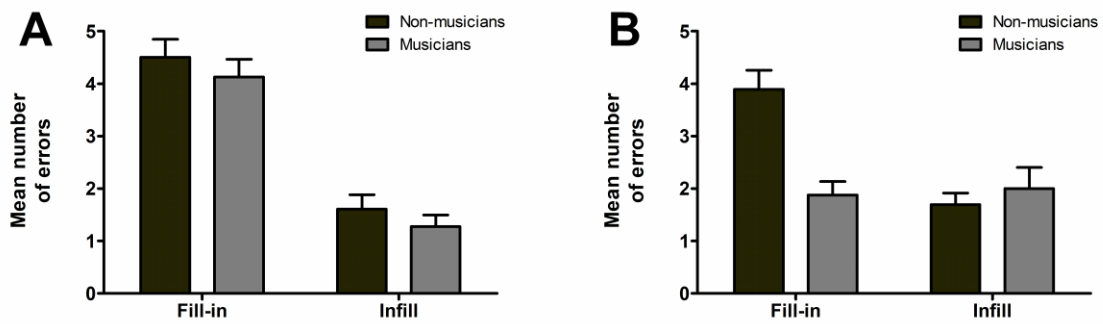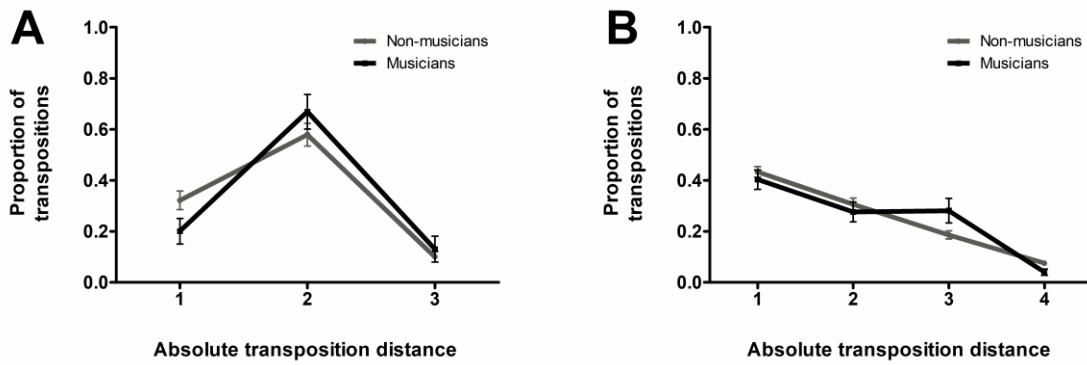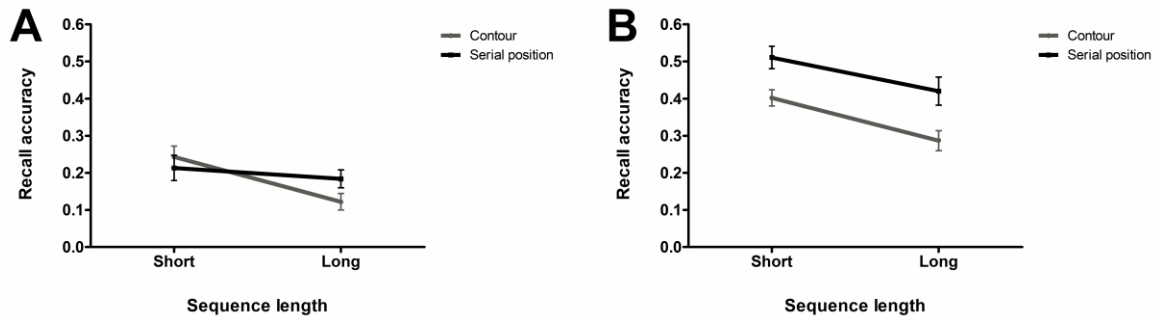
Figure 5. (A) Means and standard errors for the proportion of transpositions in the verbal order reconstruction task, as a function of displacement distance (from 1 to 7) and participant group (musicians vs. participants with no advanced musical expertise, referred to as "non-musicians"). (B) Means and standard errors for the proportion of transpositions in the musical order reconstruction task, as a function of displacement distance (from 1 to 5) and participant group (musicians vs. participants with no advanced musical expertise, referred to as "non-musicians").

Figure 6. (A) Means and standard errors for the number of fill-in and infill errors in the verbal order reconstruction task, as a function of participant group (musicians vs. participants with no advanced musical expertise, referred to as "non-musicians"). (B) Means and standard errors for the number of fill-in and infill errors in the musical order reconstruction task, as a function of participant group (musicians vs. participants with no advanced musical expertise referred, to as "non-musicians").

Figure 7. (A) Means and standard errors for the proportion of transpositions in the musical order reconstruction task for four-tone sequences, as a function of displacement distance (from 1 to 3) and participant group (musicians vs. participants with no advanced musical expertise, referred to as "non-musicians"). (B) Means and standard errors for the proportion of transpositions in the musical order reconstruction task for five-tone sequences, as a function of displacement distance (from 1 to 4) and participant group (musicians vs. participants with no advanced musical expertise, referred to as "non-musicians").

Figure 8. (A) Means and standard errors for chance-corrected response accuracy in the musical order reconstruction task in participants with no advanced musical expertise (referred to as "non-musicians"), as a function of sequence length (short vs. long) and scoring method (contour-based or serial position-based). (B) Means and standard errors for chance-corrected response accuracy in the musical order reconstruction task in musicians, as a function of sequence length (short vs. long) and scoring method (contour-based or serial position-based).

References

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An Integrated Theory of List Memory. *Journal of Memory and Language*, *38*(4), 341–380. http://doi.org/10.1006/jmla.1997.2553

Avons, S. E. (1998). Serial report and item recognition of novel visual patterns. *British Journal of Psychology*, *89*(2), 285–308. http://doi.org/10.1111/j.2044-8295.1998.tb02685.x

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation*. San Diego, CA: Academic Press.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time Constraints and Resource Sharing in Adults' Working Memory Spans. *Journal of Experimental Psychology: General*, *133*(1), 83–100. http://doi.org/10.1037/0096-3445.133.1.83

Berkowska, M., & Dalla Bella, S. (2013). Uncovering phenotypes of poor-pitch singing: the Sung Performance Battery (SPB). *Frontiers in Psychology*, *4*, 714. http://doi.org/10.3389/fpsyg.2013.00714

Berti, S., Münzer, S., Schröger, E., & Pechmann, T. (2006). Different Interference Effects in Musicians and a Control Group. *Experimental Psychology*, *53*(2), 111–116. http://doi.org/10.1027/1618-3169.53.2.111

Berz, W. L. (1995). Working Memory in Music: A Theoretical Model. *Music Perception: An Interdisciplinary Journal*, *12*(3), 353–364. http://doi.org/10.2307/40286188

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of Training between Music and Speech: Common Processing, Attention, and Memory. *Frontiers in Psychology*, *2*, 94. http://doi.org/10.3389/fpsyg.2011.00094

Bhatarah, P., Ward, G., Smith, J., & Hayes, L. (2009). Examining the relationship between free recall and immediate serial recall: Similar patterns of rehearsal and similar effects of word length, presentation rate, and articulatory suppression. *Memory & Cognition*, *37*(5), 689–713. http://doi.org/10.3758/MC.37.5.689

Botvinick, M. M., & Plaut, D. C. (2006). Short-term memory for serial order: A recurrent neural network model. *Psychological Review*, *113*(2), 201–233. http://doi.org/10.1037/0033-

295X.113.2.201

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539–576. http://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. http://doi.org/10.1037/0033-295X.107.1.127

Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language*, *31*(4), 429–460. http://doi.org/10.1016/0749-596X(92)90022-P

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*(3), 551–581. http://doi.org/10.1037/0033-295X.106.3.551

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. http://doi.org/10.1016/j.jml.2006.08.005

Chan, A. S., Ho, Y.-C., & Cheung, M.-C. (1998). Music training improves verbal memory. *Nature*, *396*(6707), 128–128. http://doi.org/10.1038/24075

Cortis, C., Dent, K., Kennett, S., & Ward, G. (2015). First things first: Similar list length and output order effects for verbal and nonverbal stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1179–1214. http://doi.org/10.1037/xlm0000086

Cowan, N. (1995). *Attention and Memory: An Integrated Framework*. Oxford, UK: Oxford University Press. http://doi.org/10.1093/acprof:oso/9780195119107.001.0001

Cowan, N., Saults, J. S., Elliott, E. M., & Moreno, M. V. (2002). Deconfounding Serial Recall. *Journal of Memory and Language*, *46*(1), 153–177. http://doi.org/10.1006/jmla.2001.2805

Dienes, Z. (2011). Bayesian Versus Orthodox Statistics: Which Side Are You On? *Perspectives on Psychological Science*, *6*(3), 274–290. http://doi.org/10.1177/1745691611406920

Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology*, *72*, 78–89. http://doi.org/10.1016/j.jmp.2015.10.003

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, *85*(4), 341–354. http://doi.org/10.1037/0033-295X.85.4.341

Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception & Psychophysics*, *50*(4), 305–313. http://doi.org/10.3758/BF03212222

Dowling, W. J., Bartlett, J. C., Halpern, A. R., & Andrews, M. W. (2008). Melody recognition at fast and slow tempos: Effects of age, experience, and familiarity. *Perception & Psychophysics*, *70*(3), 496–502. http://doi.org/10.3758/PP.70.3.496

Dowling, W. J., & Tillmann, B. (2014). Memory Improvement While Hearing Music. *Music Perception: An Interdisciplinary Journal*, *32*(1), 11–32. http://doi.org/10.1525/mp.2014.32.1.11

Edworthy, J. (1985). Interval and Contour in Melody Processing. *Music Perception: An Interdisciplinary Journal*, *2*(3), 375–388. http://doi.org/10.2307/40285305

Farrell, S., Hurlstone, M. J., & Lewandowsky, S. (2013). Sequential dependencies in recall of sequences: Filling in the blanks. *Memory & Cognition*, *41*(6), 938–952. http://doi.org/10.3758/s13421-013-0310-0

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, *9*(1), 59–79. http://doi.org/10.3758/BF03196257

Farrell, S., & Lewandowsky, S. (2004). Modelling transposition latencies: Constraints for theories of serial order memory. *Journal of Memory and Language*, *51*(1), 115–135. http://doi.org/10.1016/j.jml.2004.03.007

Franklin, M. S., Rattray, K., Sledge Moore, K., Moher, J., Yip, C.-Y., & Jonides, J. (2008). The effects of musical training on verbal memory. *Psychology of Music*, *36*(3), 353–365. http://doi.org/10.1177/0305735607086044

Gmeindl, L., Walsh, M., & Courtney, S. M. (2011). Binding serial order to representations in working memory: a spatial/verbal dissociation. *Memory & Cognition*, *39*(1), 37–46. http://doi.org/10.3758/s13421-010-0012-9

Gorin, S., Kowialiewski, B., & Majerus, S. (2016). Domain-Generality of Timing-Based Serial Order Processes in Short-Term Memory: New Insights from Musical and Verbal Domains. *PLOS ONE*, *11*(12), e0168699. http://doi.org/10.1371/journal.pone.0168699

Gorin, S., Mengal, P., & Majerus, S. (2017). Temporal grouping effects in musical short-term memory. *Memory*, 1–13. http://doi.org/10.1080/09658211.2017.1414848

Greene, R. L., & Samuel, A. G. (1986). Recency and suffix effects in serial recall of musical stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(4), 517–524. http://doi.org/10.1037/0278-7393.12.4.517

Guérard, K., & Tremblay, S. (2008). Revisiting evidence for modularity and functional equivalence across verbal and spatial domains in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(3), 556–569. http://doi.org/10.1037/0278-7393.34.3.556

Hansen, M., Wallentin, M., & Vuust, P. (2013). Working memory and musical competence of musicians and non-musicians. *Psychology of Music*, *41*(6), 779–793. http://doi.org/10.1177/0305735612452186

Hartley, T., Hurlstone, M. J., & Hitch, G. J. (2016). Effects of rhythm on memory for spoken sequences: A model and tests of its stimulus-driven mechanism. *Cognitive Psychology*, *87*, 135–178. http://doi.org/10.1016/j.cogpsych.2016.05.001

Henson, R. N. A. (1996). Short-term memory for serial order. Unpublished doctoral thesis, Cambridge University, UK. Retrieved from http://www.mrc-cbu.cam.ac.uk/people/rik.henson/personal/thesis/

Henson, R. N. A. (1998). Short-Term Memory for Serial Order: The Start-End Model. *Cognitive Psychology*, *36*(2), 73–137. http://doi.org/10.1006/cogp.1998.0685

Henson, R. N. A., Hartley, T., Burgess, N., Hitch, G. J., & Flude, B. (2003). Selective interference with verbal short-term memory for serial order information: A new paradigm and tests of a timing-signal hypothesis. *The Quarterly Journal of Experimental Psychology Section A*, *56*(8), 1307–1334. http://doi.org/10.1080/02724980244000747

Hurlstone, M. J., & Hitch, G. J. (2015). How is the serial order of a spatial sequence represented? Insights from transposition latencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 295–324. http://doi.org/10.1037/a0038223

Hurlstone, M. J., & Hitch, G. J. (2017). How Is the Serial Order of a Visual Sequence Represented? Insights From Transposition Latencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. http://doi.org/10.1037/xlm0000440

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An

overview of the literature and directions for future research. *Psychological Bulletin*, *140*(2),

339–373. http://doi.org/10.1037/a0034221

JASP Team. (2017). JASP (version 0.8.1)[Computer software].

Jefferies, E., Frankish, C. R., & Lambon Ralph, M. A. (2006). Lexical and semantic influences on

item and order memory in immediate serial recognition: Evidence from a novel task. *The*

Quarterly Journal of Experimental Psychology, *59*(5), 949–964.

http://doi.org/10.1080/02724980543000141

Johnson, A. J., Shaw, J., & Miles, C. (2016). Tactile order memory: evidence for sequence learning

phenomena found with other stimulus types. *Journal of Cognitive Psychology*, *28*(6), 718–725.

http://doi.org/10.1080/20445911.2016.1186676

Jones, D. M., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and

spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning,*

Memory, and Cognition, *21*(4), 1008–1018. http://doi.org/10.1037/0278-7393.21.4.1008

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect:

Implications for phonological coding in working memory. *Journal of Experimental Psychology:*

Learning, Memory, and Cognition, *19*(2), 369–381. http://doi.org/10.1037/0278-7393.19.2.369

Kieras, D. E., Meyer, D. E., Mueller, S., & Seymour, T. (1999). Insights into Working Memory from

the Perspective of the EPIC Architecture for Modeling Skilled Perceptual-Motor and Cognitive

Human Performance. In A. Miyake & P. Shah (Eds.), *Models of Working Memory* (pp. 183–

223). Cambridge, UK: Cambridge University Press.

http://doi.org/10.1017/CBO9781139174909.009

Kunert, R., Willems, R. M., & Hagoort, P. (2016). An Independent Psychometric Evaluation of the

PROMS Measure of Music Perception Skills. *PLOS ONE*, *11*(7), e0159103.

http://doi.org/10.1371/journal.pone.0159103

Law, L. N. C., & Zentner, M. R. (2012). Assessing Musical Abilities Objectively: Construction and

Validation of the Profile of Music Perception Skills. *PLoS ONE*, *7*(12), e52508.

http://doi.org/10.1371/journal.pone.0052508

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian Cognitive Modeling: A Pratical Course*.

Cambridge: Cambridge University Press. http://doi.org/10.1017/CBO9781139087759

Lee, Y., Lu, M., & Ko, H. (2007). Effects of skill training on working memory capacity. *Learning and Instruction*, *17*(3), 336–344. http://doi.org/10.1016/j.learninstruc.2007.02.010

Lerdahl, F., & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Lewandowsky, S., & Farrell, S. (2008). Short-Term Memory: New Data and a Model. In H. R. Brian (Ed.), *The Psychology of Learning and Motivation* (Vol. 49, pp. 1–48). San Diego, CA: Academic Press. http://doi.org/10.1016/S0079-7421(08)00001-7

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

Logie, R. H., Saito, S., Morita, A., Varma, S., & Norris, D. G. (2016). Recalling visual serial order for verbal sequences. *Memory & Cognition*, *44*(4), 590–607. http://doi.org/10.3758/s13421-015-0580-9

Majerus, S. (2013). Language repetition and short-term memory: an integrative framework. *Frontiers in Human Neuroscience*, *7*, 357. http://doi.org/10.3389/fnhum.2013.00357

Majerus, S., Cowan, N., Peters, F., Van Calster, L., Phillips, C., & Schrouff, J. (2016). Cross-Modal Decoding of Neural Patterns Associated with Working Memory: Evidence for Attention-Based Accounts of Working Memory. *Cerebral Cortex*, *26*(1), 166–179. http://doi.org/10.1093/cercor/bhu189

Majerus, S., Poncelet, M., Elsen, B., & van der Linden, M. (2006). Exploring the relationship between new word learning and short-term memory for serial order recall, item recall, and item recognition. *European Journal of Cognitive Psychology*, *18*(6), 848–873. http://doi.org/10.1080/09541440500446476

Majerus, S., Poncelet, M., Van der Linden, M., & Weekes, B. S. (2008). Lexical learning in bilingual adults: The relative importance of short-term memory for serial order and phonological knowledge. *Cognition*, *107*(2), 395–419. http://doi.org/10.1016/j.cognition.2007.10.003

Mathias, B., Pfordresher, P. Q., & Palmer, C. (2015). Context and meter enhance long-range planning in music performance. *Frontiers in Human Neuroscience*, *8*, 1040.

http://doi.org/10.3389/fnhum.2014.01040

Maybery, M. T., Parmentier, F. B. R., & Jones, D. M. (2002). Grouping of list items reflected in the

timing of recall: implications for models of serial verbal memory. *Journal of Memory and*

Language, *47*(3), 360–385. http://doi.org/10.1016/S0749-596X(02)00014-1

Mondor, T. A., & Morin, S. R. (2004). Primacy, Recency, and Suffix Effects in Auditory Short-Term

Memory for Pure Tones: Evidence From a Probe Recognition Paradigm. *Canadian Journal of*

Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, *58*(3), 206–219.

http://doi.org/10.1037/h0087445

Morey, R. D. (2015). Workshop: Bayesian analysis with BayesFactor. Katholieke Universiteit

Leuven, Belgium.

Müllensiefen, D., Gingras, B., Musil, J. J., & Stewart, L. (2014). The Musicality of Non-Musicians:

An Index for Assessing Musical Sophistication in the General Population. *PLoS ONE*, *9*(2),

e89642. http://doi.org/10.1371/journal.pone.0089642

Oberauer, K. (2003). Understanding serial position curves in short-term recognition and recall.

Journal of Memory and Language, *49*(4), 469–483. http://doi.org/10.1016/S0749-

596X(03)00080-9

Ockelford, A. (2007). A Music Module in Working Memory? Evidence from the Performance of a

Prodigious Musical Savant. *Musicae Scientiae*, *11*(2 suppl), 5–36.

http://doi.org/10.1177/10298649070110S202

Page, M. P. A., & Norris, D. G. (1998). The primacy model: A new model of immediate serial recall.

Psychological Review, *105*(4), 761–781. http://doi.org/10.1037/0033-295X.105.4.761-781

Page, M. P. A., & Norris, D. G. (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. http://doi.org/10.1098/rstb.2009.0173

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician Enhancement for Speech-In-

Noise. *Ear and Hearing*, *30*(6), 653–661. http://doi.org/10.1097/AUD.0b013e3181b412e9

Parmentier, F. B. R., Andrés, P., Elford, G., & Jones, D. M. (2006). Organization of visuo-spatial

serial memory: interaction of temporal order with spatial and temporal grouping. *Psychological*

Research Psychologische Forschung, *70*(3), 200–217. http://doi.org/10.1007/s00426-004-0212-7

Parmentier, F. B. R., King, S., & Dennis, I. (2006). Local temporal distinctiveness does not benefit auditory verbal and spatial serial recall. *Psychonomic Bulletin & Review*, *13*(3), 458–465. http://doi.org/10.3758/BF03193870

Parmentier, F. B. R., Maybery, M. T., & Jones, D. M. (2004). Temporal grouping in auditory spatial serial memory. *Psychonomic Bulletin & Review*, *11*(3), 501–507. http://doi.org/10.3758/BF03196602

Pechmann, T., & Mohr, G. (1992). Interference in memory for tonal pitch: Implications for a working-memory model. *Memory & Cognition*, *20*(3), 314–320. http://doi.org/10.3758/BF03199668

Pfordresher, P. Q., & Brown, S. (2007). Poor-Pitch Singing in the Absence of "Tone Deafness." *Music Perception: An Interdisciplinary Journal*, *25*(2), 95–115. http://doi.org/10.1525/mp.2007.25.2.95

Pfordresher, P. Q., Brown, S., Meier, K. M., Belyk, M., & Liotti, M. (2010). Imprecise singing is widespread. *The Journal of the Acoustical Society of America*, *128*(4), 2182–2190. http://doi.org/10.1121/1.3478782

Pfordresher, P. Q., Palmer, C., & Jungers, M. K. (2007). Speed, Accuracy, and Serial Order in Sequence Production. *Cognitive Science*, *31*(1), 63–98. http://doi.org/10.1080/03640210709336985

Plancher, G., Lévêque, Y., Fanuel, L., Piquandet, G., & Tillmann, B. (2017). Boosting Maintenance in Working Memory With Temporal Regularities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. http://doi.org/10.1037/xlm0000481

Posner, M. I. (1964). Rate of presentation and order of recall in immediate memory. *British Journal of Psychology*, *55*(3), 303–306. http://doi.org/10.1111/j.2044-8295.1964.tb00914.x

Raven, J. C. (1938). *Progressive Matrices: A perceptual test of intelligence*. Oxford, UK: Oxford Psychologists Press.

Rouder, J. N., Engelhardt, C. R., McCabe, S., & Morey, R. D. (2016). Model comparison in ANOVA.

Psychonomic Bulletin & Review, *23*(6), 1779–1786. http://doi.org/10.3758/s13423-016-1026-5

Ryan, J. (1969). Grouping and short-term memory: Different means and patterns of grouping. *Quarterly Journal of Experimental Psychology*, *21*(2), 137–147. http://doi.org/10.1080/14640746908400206

Saito, S. (2001). The phonological loop and memory for rhythms: An individual differences approach. *Memory*, *9*(4–6), 313–322. http://doi.org/10.1080/09658210143000164

Saito, S., Logie, R. H., Morita, A., & Law, A. (2008). Visual and phonological similarity effects in verbal immediate serial recall: A test with kanji materials. *Journal of Memory and Language*, *59*(1), 1–17. http://doi.org/10.1016/j.jml.2008.01.004

Schendel, Z. A., & Palmer, C. (2007). Suppression effects on musical and verbal memory. *Memory & Cognition*, *35*(4), 640–650. http://doi.org/10.3758/BF03193302

Schulze, K., Dowling, W. J., & Tillmann, B. (2012). Working Memory for Tonal and Atonal Sequences during a Forward and a Backward Recognition Task. *Music Perception: An Interdisciplinary Journal*, *29*(3), 255–267. http://doi.org/10.1525/mp.2012.29.3.255

Schulze, K., Mueller, K., & Koelsch, S. (2011). Neural correlates of strategy use during auditory working memory in musicians and non-musicians. *European Journal of Neuroscience*, *33*(1), 189–196. http://doi.org/10.1111/j.1460-9568.2010.07470.x

Smyth, M. M., Hay, D. C., Hitch, G. J., & Horton, N. J. (2005). Serial position memory in the visual—spatial domain: Reconstructing sequences of unfamiliar faces. *The Quarterly Journal of Experimental Psychology Section A*, *58*(5), 909–930. http://doi.org/10.1080/02724980443000412

Strait, D. L., Kraus, N., Parbery-Clark, A., & Ashley, R. (2010). Musical experience shapes top-down auditory mechanisms: Evidence from masking and auditory attention performance. *Hearing Research*, *261*(1–2), 22–29. http://doi.org/10.1016/j.heares.2009.12.021

Surprenant, A. M., Kelley, M. R., Farley, L. A., & Neath, I. (2005). Fill-in and infill errors in order memory. *Memory*, *13*(3–4), 267–273. http://doi.org/10.1080/09658210344000396

Tan, L., & Ward, G. (2008). Rehearsal in immediate serial recall. *Psychonomic Bulletin & Review*, *15*(3), 535–542. http://doi.org/10.3758/PBR.15.3.535

Tremblay, S., Parmentier, F. B. R., Guérard, K., Nicholls, A. P., & Jones, D. M. (2006). A spatial

modality effect in serial memory. *Journal of Experimental Psychology: Learning, Memory, and*

Cognition, *32*(5), 1208–1215. http://doi.org/10.1037/0278-7393.32.5.1208

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems ofp values. *Psychonomic*

Bulletin & Review, *14*(5), 779–804. http://doi.org/10.3758/BF03194105

Wagenmakers, E.-J., Lee, M. D., Lodewyckx, T., & Iverson, G. J. (2008). Bayesian Versus

Frequentist Inference. In H. Hoijtink, I. Klugkist, & P. A. Boelen (Eds.), *Bayesian Evaluation of*

Informative Hypotheses (pp. 181–207). New York, NY: Springer New York.

http://doi.org/10.1007/978-0-387-09612-4_9

Ward, G., Avons, S. E., & Melling, L. (2005). Serial position curves in short-term memory:

Functional equivalence across modalities. *Memory*, *13*(3–4), 308–317.

http://doi.org/10.1080/09658210344000279

Ward, G., Tan, L., & Grenfell-Essam, R. (2010). Examining the relationship between free recall and

immediate serial recall: The effects of list length and output order. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, *36*(5), 1207–1241.

http://doi.org/10.1037/a0020122

Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term

memory for verbal and musical sequences: Comparing phonological similarity and pitch

proximity. *Memory & Cognition*, *38*(2), 163–175. http://doi.org/10.3758/MC.38.2.163

Williamson, V. J., Mitchell, T., Hitch, G. J., & Baddeley, A. D. (2010). Musicians' memory for verbal

and tonal materials under conditions of irrelevant sound. *Psychology of Music*, *38*(3), 331–350.

http://doi.org/10.1177/0305735609351918