Original Article

QJEP

Quarterly Journal of Experimental Psychology
2018, Vol. 71(6) 1367–1381

DOI: 10.1177/1747021817734978
qjep.sagepub.com
SAGE

Memory for melodies in unfamiliar tuning systems: Investigating effects of recency and number of intervening items

Steffen A Herff[1], Kirk N Olsen[1,2], Roger T Dean[1] and Jon Prince[3]

Abstract

In a continuous recognition paradigm, most stimuli elicit superior recognition performance when the item to be recognized is the most recent stimulus (a recency-in-memory effect). Furthermore, increasing the number of intervening items cumulatively disrupts memory in most domains. Memory for melodies composed in familiar tuning systems also shows superior recognition for the most recent melody, but no disruptive effects from the number of intervening melodies. A possible explanation has been offered in a novel regenerative multiple representations (RMR) conjecture. The RMR assumes that prior knowledge informs perception and perception influences memory representations. It postulates that melodies are perceived, thus also represented, simultaneously as integrated entities and also as their components (such as pitches, pitch intervals, short phrases and rhythm). Multiple representations of the melody components and melody as a whole can restore one another, thus providing resilience against disruptive effects from intervening items. The conjecture predicts that melodies in an unfamiliar tuning system are not perceived as integrated melodies and should (a) disrupt recency-in-memory advantages and (b) facilitate disruptive effects from the number of intervening items. We test these two predictions in three experiments. Experiments 1 and 2 show that no recency-in-memory effects emerge for melodies in an unfamiliar tuning system. In Experiment 3, disruptive effects occurred as the number of intervening items and unfamiliarity of the stimuli increased. Overall, results are coherent with the predictions of the RMR conjecture. Further investigation of the conjecture's predictions may lead to greater understanding of the fundamental relationships between memory, perception and behavior.

Keywords

Memory; interference; music perception; recognition; unfamiliar tuning systems

Received: 7 October 2016; revised: 13 February 2017; accepted: 17 April 2017

Recognition performance in most domains decreases as the number of intervening items increases between the first and second occurrences of a target stimulus. This phenomenon has been subject to much experimental investigation in a variety of tasks using a variety of stimuli, such as digits, letter trigrams, word lists and pairs, and faces (Bui, Maddox, Zou, & Hale, 2014; Donaldson & Murdock, 1968; Hockley, 1992; Konkle, Brady, Alvarez, & Oliva, 2010; Olson, 1969; Poon & Fozard, 1980; Rakover & Cahlon, 2001; Sadeh, Ozubko, Winocur, & Moscovitch, 2014). Despite the widespread nature of this effect, increasing the number of intervening items does not always lead to decrements in recognition performance. In the visual domain, for example, disruptive effects from the number of intervening items occur for pictures of everyday objects, but not for line drawings of everyday objects (Berman, Friedman, & Cramer, 1991; Friedman, 1990; Konkle et al., 2010). In the auditory domain, disruptive effects are apparent in spoken words and recognition of single musical notes, but not for recognition of

[1]The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Penrith, NSW, Australia
[2]Department of Psychology, Macquarie University, North Ryde, NSW, Australia
[3]School of Psychology and Exercise Science, Murdoch University, Murdoch, WA, Australia

Corresponding author:
Steffen A Herff, The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia.
Email: s.herff@westernsydney.edu.au

novel melodies in familiar musical tuning systems such as 12-Tone-Equal-Temperament (12-TET) (Buchsbaum, Padmanabhan, & Berman, 2011; Campeanu, Craik, Backer, & Alain, 2014; D. Deutsch, 1970, 1975; Herff, Olsen, & Dean, submitted).

For example, Schellenberg and Habashi (2015) investigated how recognition memory for melodies in a familiar tuning system decays over time and found delays of up to 1 week had minimal effects. Complementing Schellenberg and Habashi's results on decay, Herff et al. (submitted) recently investigated interference and reported no disruptive effects from the number of intervening items on melody recognition using up to 197 intervening melodies. Herff et al. (submitted) did, however, find a melody recognition advantage for immediate repetition. Recognition advantages for immediate repetition, or recency effects, are commonly observed in memory literature as a memory advantage for the last stimulus presented (Berz, 1995; Dowling, 1973; Greene & Samuel, 1986; Jahnke, 1963; Roberts, 1986). However, melodies present a special case for recency because they consist of a sequence of notes. In this case, the last stimulus encountered is always the final note, rather than the last melody. To trigger a recency effect, melodies would need to be perceived as perceptually integrated entities, rather than just individual notes. Previous literature suggests that melodies are indeed perceived as coherent integrated wholes (D. Deutsch, 1986; Dowling, 1991; Krumhansl, 1991). Herff et al. (submitted) have proposed a novel regenerative multiple representations (RMR) conjecture to explain why memory for melody not only exhibits recency-in-memory effects but is also not cumulatively disrupted as the number of intervening melodies increases.

RMR in memory for melody

The RMR conjecture combines and generalizes well-established memory theories. The conjecture assumes, first, that previous knowledge or experience directly influences our perception and, second, that perception determines the formation of future memories. In other words, we learn the most relevant way to perceive our environment, perceive objects according to this information and form memories according to the perception. This is analogous to Deutsch's concept of *Pertinence* (D. Deutsch, 1986; see also J. A. Deutsch and Deutsch, 1965). Pertinence is described as a weighting of perception in the form of awareness based on the current situation and prior knowledge. We apply this term to memory, inferring that previously learned perceptual relevance directly influences perception itself and therefore influences subsequent formation of new memories based on these perceptions.

Following the assumption that experience directly influences perception and that perception determines the formation of future memories, the conjecture postulates the formation of multiple memory representations in cases where prior knowledge dictates multiple perceptual relevancies. This is, for example, demonstrated in Stroop effects where the written name of a color and the actual color of a word mutually interfere (Stroop, 1992). This phenomenon would not occur if the word was written in an unfamiliar language. In this case, prior knowledge cannot inform perceptual integration of the letters in the words into a meaningful word.

The third assumption of the RMR conjecture is a common one: memory representations are subject to decay and interference. Memory models often implement this process mathematically in the form of a decay parameter, an interference parameter or both (Norman, 2013; Oberauer & Lewandowsky, 2011; Oberauer, Lewandowsky, Farrell, Jarrold, & Greaves, 2012).

The fourth assumption of the RMR conjecture is analogous to—and a generalized form of—Paivio's dual-coding theory (Paivio, 1969). In the context of words and images, the dual-coding theory describes exactly two representations that assist each other in retrieval, therefore increasing the chance of remembering a stimulus. The RMR conjecture generalizes this case and postulates any possible number of representations that are not necessarily just words and images. Prior knowledge determines the number and nature of representations. This is important for a theory applied to music, where multiple representations beyond two are highly likely due to music's perceptual reliance on multiple levels, such as its underlying parts (e.g., notes, intervals, phrases), integrations of the underlying parts (e.g., coherent melody), socio-cognitive components (e.g., emotions) and motor responses evoked by its temporal organization (Barascud, Pearce, Griffiths, Friston, & Chait, 2016; D. Deutsch, 1986; Krumhansl, 1991, p. 295; Schneider, 1997, p. 119; Zatorre, Chen, & Penhune, 2007). Analogous to Paivio's dual-coding theory, RMR assumes that multiple representations can regenerate each other. This would mean that once encoded, such RMR are robust to the interference of time (decay) and intervening items since multiple representations would be theoretically required to drop below a hypothetical retrieval threshold before retrieval is impossible. For example, to truly forget a melody beyond recognition, the RMR conjecture assumes that not only the representation of the integrated melody but also the representations of all melody-specific underlying components, such as short phrases within the melody, need to be inaccessible. A promising candidate mechanism for memory regeneration in the context of music is melodic expectancy. Due to familiarity with the underlying rules of a music tradition, listeners form strong melodic expectancies. These expectancies can be used to predict what comes next in a melody (Margulis, 2005; Pearce, 2014; Schellenberg, 1996) and, theoretically, could also be used to interpolate forgotten parts of a melody.

RMR: implications for unfamiliar tuning systems

Integrating notes, intervals and short phrases into a coherent melody requires familiarity with the underlying rules of how individual components inter-relate (Cui, Collett, Troje, & Cuddy, 2015; Saffran, Johnson, Aslin, & Newport, 1999; Schon & Francois, 2011; Tillmann & McAdams, 2004). The regeneration process described above reduces the disruptive effect of the number of intervening items on melody recognition and requires multiple representations of the same melody: representations that code—at least partially—overlapping features of the stimulus. For example, the representations of a melody's components (e.g., notes, intervals, short phrases) as well as their integrated whole (the melody itself) can have redundancy. To form a representation of the integrated melody, the melodic sequence first needs to be perceived as an integrated melody. This means that disrupting the perception of an integrated melody should simultaneously disrupt the formation of its memory representation. As a result, the regeneration process cannot be fully triggered. In this case, the RMR conjecture predicts stronger disruptive effects from the number of intervening items because listeners should only perceive and form representations of the melody's underlying components, and not of an integrated melody as a whole. Melodies that are sounded in an unfamiliar tuning system are a useful context to test this hypothesis. From classical to modern pop music, the tuning system most Western listeners are familiar with is 12-TET—a tuning system that describes a division of a musical octave in 12 equal step sizes. However, there is an infinite number of other possible tuning systems.

As stated above, listeners presented with melodies in an unfamiliar tuning system should perceive and form a representation of a melody's underlying components. However, assuming the listener is unfamiliar with that tuning system's rules and relations between the components required to integrate the melodies into a coherent whole, they should not be capable of perceiving large-scale melodic structure (e.g., Castellano, Bharucha, & Krumhansl, 1984), thus not forming a representation of a melody as a whole. This should affect the regeneration process, resulting in a stronger disruptive effect from the number of intervening melodies on recognition of melodies composed in an unfamiliar tuning system, relative to melodies composed in a familiar tuning system. In other words, recognition performance in response to melodies in an unfamiliar tuning system should decline as the number of intervening melodies increases.

Previous research has observed recency-in-memory effects in recognition of melodies in the familiar 12-TET tuning system (Berz, 1995; Dowling, 1973; Herff et al., submitted). If a melody in a familiar tuning system presented at test was also the previously presented melody, then a recognition advantage emerges relative to when the tested melody was not the previously presented stimulus. This recency-in-memory effect suggests that melodies in a familiar tuning system are indeed treated as integrated wholes. This is because the recency-in-memory effect is defined as enhanced recognition for the last stimulus presented—in this case, the melody as a whole. If a melodic sequence is not perceived as an integrated melody, then the "last stimulus encountered" is no longer the melody as a whole; rather, it is most likely the last note or at most the last phrase of the melody. As a result, no recognition enhancement for the most recent melody would be expected because the most recent stimulus is no longer defined at the level of the melody as a whole.

If unfamiliar tuning systems disrupt integration of melodies as a whole (as suggested by the RMR conjecture), then no recency-in-memory effect should occur. This is because recognition memory in this context can only rely on representations of the melody's underlying components, rather than multiple representations that include the melody's underlying components *and* the melody as a coherent whole. Furthermore, cumulative disruptive effects of the number of intervening items should also occur. These hypotheses will be investigated here.

Aim, design and hypothesis

The aim of this study is to investigate recency-in-memory and cumulative disruptive effects from an increase in the number of intervening melodies in memory for melodies in an unfamiliar tuning system. The study closely follows the design of previous work investigating these effects using stimuli composed in a familiar tuning system (Herff et al., submitted). Here, Experiment 1 tests the early time-course of recognition of melodies in an unfamiliar tuning system using an explicit memory task in a continuous recognition paradigm and varying the number of intervening melodies between immediate repetition (0) and 13. Experiment 2 provides a close replication of Experiment 1, however with an indirect memory task (perceived familiarity). Such a task was chosen because there is evidence that the level of task awareness influences some memory for melody phenomena (Gaudreau & Peretz, 1999; Halpern & Bartlett, 2010; Halpern & O'Connor, 2000). Experiment 3 tests whether there are cumulative effects from the number of intervening melodies on melody recognition using up to 107 intervening melodies and two repetitions of each melody.

Experiment 1: 0-13 intervening melodies in an unfamiliar tuning system

Method

Participants. In total, 37 undergraduate students were recruited from Western Sydney University (M_{age} = 22 years,

SD_{age} = 5.5; 7 males, 30 females). Participants were required to have had less than 2 years of musical training and no hearing impairments. Participants were reimbursed with course credit as part of university course requirements.

Stimuli and equipment. Stimuli from all experiments can be found in Supplement S1 Stimuli.zip. The aim of the stimulus generation process was to generate melodies within an unfamiliar tuning system. To achieve this, a set of tonal melodies that followed the Western-tonal tradition were chosen and placed in a specially designed tuning system comprising carefully defined similarities and dissimilarities to the familiar 12-TET. The original 12-TET stimulus set consisted of a mathematically and perceptually tested subset of European folk songs that has been used in a previous related study (Herff et al., submitted). Previously, a pilot study established the response pattern of a group of listeners to unfamiliar melodies and removed any of the original stimuli that showed atypical familiarity response distributions (details in Herff et al. (submitted)). Here, we used the same stimuli but retuned to a new unfamiliar tuning system, as described below.

The construction of the unfamiliar tuning system was based on the following principles: First, to ensure that potential effects from the number of intervening items can be interpreted as arising from the unfamiliarity of the tuning system, rather than contaminated by perceptual difficulties, the smallest possible step in this unfamiliar tuning system was made large enough to ensure that two pitches are easily discriminated. Pitch interval sizes are commonly measured in log-frequency units such as *cents* because pitch, particularly in a musical context, is more linearly related to log-frequency than it is to frequency. The 12-TET has step sizes of 100 cents (the term *step* is here used for the interval between adjacent pitches in the tuning system). Previous research suggests that pitches separated by 40 cents or more can be discriminated approximately, as well as those separated by the familiar 100 cents (Parncutt & Cohen, 1995).

Second, an equally tempered tuning system was sought, but we did not impose any requirement for our novel tuning systems to contain octaves. Typically, pitches an octave apart (double or half the frequency) are heard as similar (if not identical) and often have the same musical function (D. Deutsch & Boulanger, 1984; Krumhansl & Shepard, 1979). One advantage of not privileging octaves in our tuning system is that a greater number of equally tempered tuning systems can be considered in the following step. In summary, we aimed to develop an equally tempered tuning system that minimized perceived similarity to 12-TET using step sizes between 40 and 100 cents and without prioritization of octave relationships because the RMR conjecture predicts larger effects the more unfamiliar a tuning system is.

Dissimilarity between tunings systems was assessed using a psychoacoustically informed and perceptually tested model of tonal affinity and similarity developed by Milne and colleagues, which was first described in Milne, Sethares, Laney, and Sharp (2011) and elaborated in Milne (2013). These papers present a related family of models of tonal similarity/affinity, one of which (*relative dyad expectation vectors*) can assess the similarity of the interval content of one scale or tuning system with the interval content of another while taking into account uncertainties of pitch perception.

The model has one free parameter *sigma*, which is the standard deviation (SD) of a Gaussian smoothing kernel applied to the log-frequencies of each scale tone. Sigma models the inaccuracy of pitch perception, and a value of 6 cents was chosen because sigma has optimized to approximately 6 cents when fitted to data from three different experiments (Milne & Holland, 2016; Milne, Laney, & Sharp, 2015, 2016). When both tuning systems are smoothed in this way, their similarity is determined by their cosine similarity, which is 0 when the two systems are orthogonal (maximally dissimilar) and 1 when they are identical.

To find the equally tempered tuning that is maximally dissimilar from 12-TET, a MATLAB routine generated equal temperament tuning systems with steps between 40 and 100 cents, using an increment of 0.01 cents (so the first equally tempered tuning system has steps all of size 40 cents, the second system has steps all of size 40.01 cents, the third has steps all of size 40.02 cents and so on until the final system has steps of 100 cents). For each tuning system, the routine calculated the cosine similarity with the 12-TET system. The tuning system with the highest dissimilarity was chosen. This maximally dissimilar equally tempered tuning had steps of size 88.08 cents and is denoted as 88.08-CET (*cents equal temperament*). The new 88.08-CET shows a cosine similarity of 0.28058 with the familiar 12-TET system, as measured by the tonal affinity model (Milne, 2013; Milne et al., 2011).

The original melodies were played through the physical synthesis PianoTeq 5 using the 88.08-CET system. The Scala file of the 88.08-CET system can be found in Supplement S1 Stimuli.zip/Experiment1And2/8808cET.scl. The step numbers between notes were held constant, mapping the original 12-TET melodies to the new 88.08-CET tuning system. That means that if in the original melodies a note was two step sizes apart from an adjacent note in the sequence, it will still be two step sizes apart from the surrounding notes in the new tuning system, although the size of the steps has changed. The result was 50 novel melodies in an unfamiliar tuning system whose rhythms and contours resemble the Western-tonal melodies that previously exhibited recency-in-memory effects on recognition for melodies, yet no disruptive effects from the number of intervening melodies (Herff et al., submitted). The melodies had a mean duration of 10.80 s ($SD=2.29$) and consisted of 15-78 notes. An analysis of the melodies can be found in Supplement S1 Stimuli.zip/ Experiment1And2/ MusicalFeatures.csv, which describes every melody along with 21 musical features.[1] Melodies were presented diotically in stereo through Sennheiser HD 25 headphones using an UA-25 Edirol external USB soundcard.

Procedure. Testing took place in a sound-attenuated booth provided by the MARCS Institute, Sydney, Australia. The procedure closely followed Herff et al. (submitted). Participants provided informed consent, and demographic questionnaires were administered. Standardized instructions appeared on a computer screen and participants were informed that they would hear "*many different melodies one after another*" and that they were required to indicate whether "*a melody has already been presented in this experiment*." Melodies were presented in random order in a continuous recognition paradigm (Shepard & Teghtsoonian, 1961). In this task, melodies are continuously presented and a response is required after each melody. Participants responded using the mouse to click one of two buttons that appeared on the screen after the presentation of each melody. One button was labeled "*New*" and the other one "*Old*." The "*New*" button was to be clicked if the first presentation of a melody was perceived, and the "*Old*" button was clicked if a melody was perceived as previously presented. Each new trial was initiated as soon as a participant gave a response. Unknown to the participant, the number of melodies intervening between the two presentations of a given melody varied between 0, 1, 2, 3, 4, 7 and 13. The order of melodies and conditions was randomized. Participants completed six practice trials, in which they adjusted the volume to their personal preference. Testing of the 100 trials took approximately 30 min.

Statistical approach. The statistical approach was analogous to previous approaches outlined in Herff et al. (submitted). Specifically, we used generalized linear mixed-effects models to investigate the influence of the number of intervening melodies on binary melody recognition data (Experiments 1 and 3) (Baayen, 2008; Baayen, Davidson, & Bates, 2008; Judd, Westfall, & Kenny, 2012; Kass & Raftery, 1995; Kruschke, 2010, 2013; Nathoo & Masson, 2016). Furthermore, we used linear mixed-effects models to analyze continuous familiarity data (Experiment 2). The models were implemented in the R software platform (R Core Team, 2013) using the lme4 package (Bates, Maechler, Bolker, & Walker, 2013). The models consisted of the experimental fixed factor *Number of Intervening Melodies*. Random effects on *Participant* and *Melody* intercepts were included in the models (Barr, Levy, Scheepers, & Tily, 2013). Coefficient *p*-values are reported as calculated by lme4 for generalized mixed-effects models and Kenward–Roger corrected for linear mixed-effects models (Halekoh & Højsgaard, 2014a, 2014b; Kenward & Roger, 1997). Significant coefficients were also further assessed in the form of a model comparison approach (Kruschke, 2011). Each model with a significant predictor was also compared with the same model but without the significant predictor using likelihood-ratio tests (Wilks, 1938). To ensure that significance was not due to an increase in model complexity, differences in Bayesian information criterion (BIC; Schwarz, 1978) are reported in the form of ΔBIC (Kass & Raftery, 1995). A ΔBIC of 2 or greater is considered "positive" evidence in favor of the model with lower BIC. A ΔBIC difference of 6 or greater is considered "strong" evidence (Kass & Raftery, 1995).

The possibility of "response tendency shifts" throughout the course of recognition experiments is a significant issue that we address in the present set of analyses (Berch, 1976; Donaldson & Murdock, 1968; Snodgrass & Corwin, 1988). These shifts describe changes in the response bias as an experiment progresses, for example, changes in response tendencies due to fatigue. Similar to previous studies (Herff et al., submitted; Herff, Olsen, Dean, & Prince, in press), we trained participant-wise generalized mixed-effects models on "old" responses on first presentations (false alarm rates) based on trial number. The fitted model was then used to predict the probability of pressing "old" on a repetition trial, based solely on trial number. These predictions were then implemented as a fixed *Dynamic Response Tendency* factor in all models to account for individual participant response tendencies and for how these tendencies might change over the course of the experiment. Mixed-effects models also assess whether overall performance in each experiment was at chance. Consequently, coefficient *Z*-scores and *p*-values of the increase in "old" responses between first and second melody presentations are reported at the beginning of each results section.

Results

Figure 1 shows melody- and participant-wise performance. Overall, participants performed significantly above chance ($Z=15.67$, $p<0.0001$). A generalized mixed-effects model was constructed to investigate the first hypothesis that melodies in an unfamiliar tuning do not exhibit a recency-in-memory effect. This hypothesis was supported. Coefficient estimation for each number of intervening melodies showed that no number of intervening items produced significantly worse recognition performance than immediate repetition (all $p>0.10$). Furthermore, a model predicting "old" responses on melody repetitions using a random intercept for *Subject* and *Melody*, as well as a systematic factor for *Dynamic Response Tendency* ($BIC=2262.9$, $LogLik=-1116.4$), did not significantly improve when provided with the *Number of Intervening Items* ($BIC=2267.8$, $LogLik=-1115.1$, $p=0.103$).

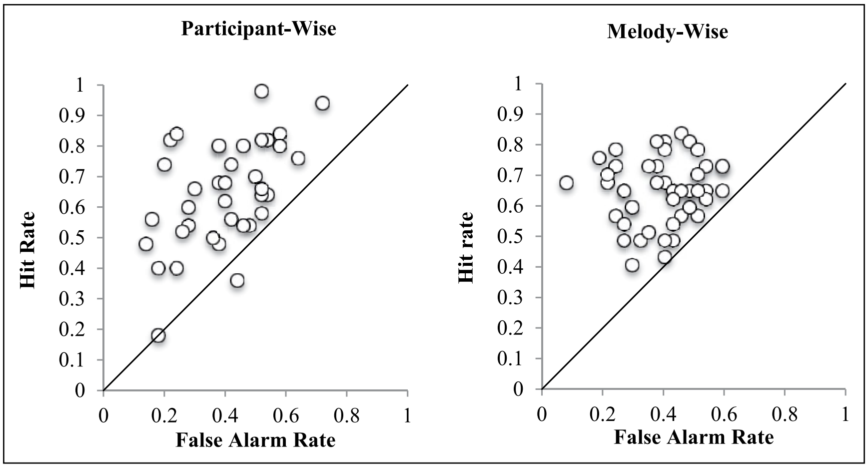

Figure 1. Hit rates and false alarm rates in Experiment 1. The left panel shows the data participant-wise and the right panel melody-wise. The reference line represents chance level.

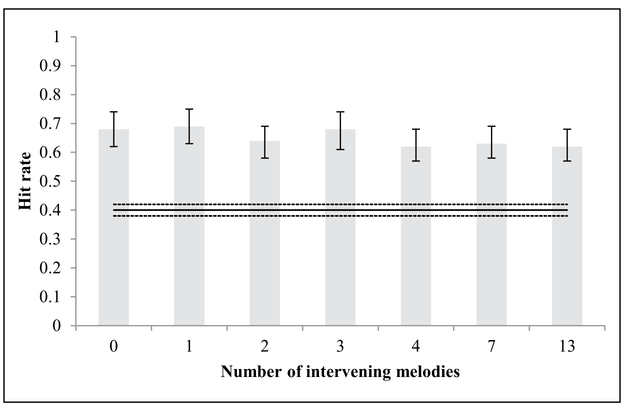

Figure 2. Mean hit rates for all seven conditions of intervening melodies in Experiment 1. Error bars show 95% confidence intervals. All conditions were statistically identical. The dark line depicts mean false alarm rates. The two dotted lines represent a 95% confidence interval around the false alarms. The difference between the false alarm rates and the hit rate can be interpreted as bias-corrected performance.

Within the tested maximum of 13 intervening melodies, the non-significant result from the intervening melodies does not support the second hypothesis that, in contrast to Western-tonal melodies, melodies in an unfamiliar tuning system show disruptive effects of the number of intervening items. Figure 2 depicts recognition performance in response to each intervening item condition.

Discussion

Experiment 1 aimed to test two hypotheses generated by the RMR conjecture. First, the conjecture predicts that melodies in an unfamiliar tuning system do not show recency-in-memory effects, unlike melodies in a familiar tuning system. The present results supported this hypothesis. Previously, participants showed recognition advantages for melodies in a familiar tuning system when the test item was the same melody as the previously presented item (Herff et al., submitted)—in other words, no intervening melodies between the first and second occurrences. The melodies presented here in an unfamiliar tuning system did not produce such recency-in-memory effects. The second hypothesis predicted that melodies in an unfamiliar tuning system show cumulative disruptive effects from the number of intervening melodies, whereas melodies in familiar tuning systems do not (cf., Herff et al., submitted). This hypothesis was not supported here using 13 intervening melodies. Specifically, the number of intervening melodies did not have a disruptive effect when composed in the unfamiliar 88.08-CET system. This result is somewhat surprising, but the effect may be due to the small number of intervening melodies (up to 13).

Other auditory stimuli such as words show disruptive effects within this range of intervening items (Buchsbaum et al., 2011). However, the melodic sequences used here are still similar to 12-TET in regard to pitch contour, rhythm and the usage of an equally tempered tuning system. This similarity could account for some of the resilience in memory that has been previously observed using Western-tonal melodies. Nevertheless, observing disruptive effects when using a greater number of intervening melodies and less familiar melodic material than Experiment 1 would provide a more comprehensive test of the RMR conjecture. We will address these issues in Experiment 3.

Alternatively, the findings here could be specific to the explicit recognition task. Previous literature suggests that the level of task awareness might influence memory for melody (Gaudreau & Peretz, 1999; Halpern & Bartlett, 2010; Halpern & O'Connor, 2000). Participants in Experiment 1 were fully aware that their memory was being tested. It is possible that the here-observed disruption of recency-in-memory effects is specific for explicit recognition and does not generalize to indirect measurements of memory. Therefore, Experiment 2 investigates the influence of intervening melodies on perceived familiarity without informing participants about the recurrent nature of the stimuli.

Experiment 2: perceived familiarity using 0-13 intervening melodies

Rather than explicit recognition performance, Experiment 2 utilized an indirect memory paradigm: perceived familiarity. Although not directly instructed to perform a memory task, participants' perceived familiarity ratings tend to be higher on the second presentation of a melody compared to the first (Herff et al., submitted). An increase in perceived familiarity can be used as a proxy of the strength of memory. Based on the results of Experiment 1 and the RMR conjecture, no recency-in-memory effect should emerge from ratings of perceived familiarity. That is, a significantly larger increase in familiarity after one intervening item, compared to all other intervening items, is not explained by a linear decrease with an overall increase in the number of intervening melodies. This hypothesis is investigated here.

Method

Participants. In total, 27 undergraduate students were recruited from the Western Sydney University (M_{age} = 24 years, SD_{age} = 10.8, 5 males/22 females) and did not participate in Experiment 1. Participants were required to have had less than 2 years of musical training and no hearing impairments. Participants were reimbursed with course credit as part of university course requirements.

Stimulus. Stimuli were identical to Experiment 1.

Procedure. The procedure closely followed that of Experiment 1. However, participants in Experiment 2 were not instructed that they were engaging in a memory task and not informed that melodies were repeated throughout the experiment. Instead, participants were prompted to indicate "*how familiar you perceive each melody to be*." Responses were made using the mouse and a vertical 100-point visual analogue scale with a spatial extent of 10 cm. The scale was labeled "familiar" at the top and "unfamiliar" at the bottom.

Results

Figure 3 shows melody- and participant-wise perceived familiarity on first and second occurrences of melodies. Coefficient estimation revealed that participants' perceived familiarity ratings were significantly higher (t = 10.69, p < 0.0001) on second presentations (M = 54.20, standard error [SE] = 5.85) than on first (M = 44.16, SE = 5.59).

A linear mixed-effects model investigated whether the previously observed disruption of a recency-in-memory effect with melodies in an unfamiliar tuning system also manifests using an indirect memory task. Disruption of a recency-in-memory effect was indeed observed. Coefficient assessment showed that the number of intervening melodies did not elicit statistically different changes in perceived familiarity (all p > 0.10). Figure 4 shows the mean familiarity ratings for each number of intervening melodies. A model predicting the change in perceived familiarity between the first and second occurrences of the melodies using a random intercept for *Subject* and *Melody* (BIC = 12,937, $LogLik$ = −6453.9) did not

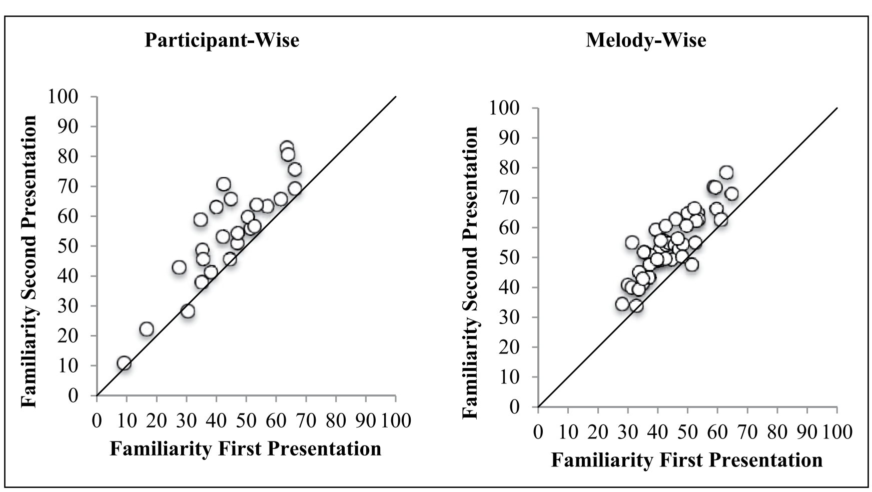

Figure 3. Perceived familiarity raw data for Experiment 2. The left panel shows participant-wise differences in perceived familiarity between the first and second occurrences of a melody. The right panel shows melody-wise differences. Overall, second presentations of melodies elicited significantly higher familiarity ratings than first presentations.

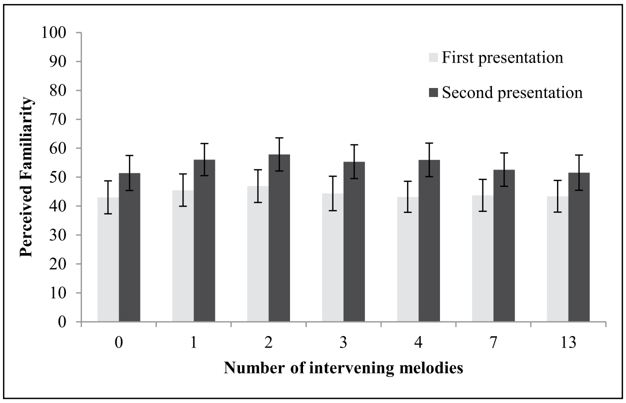

Figure 4. Perceived familiarity raw data for Experiment 2. The light bars show perceived familiarity on the first presentations of melodies, and the dark bars show perceived familiarity of the second presentations of the associated melodies. Error bars represent the standard error. Familiarity was significantly higher on second presentations. However, there were no statistically significant differences between the different conditions of intervening melodies.

improve significantly when provided with the *Number of Intervening Items* ($BIC = 12{,}942$, $LogLik = -6453.2$, $p = 0.225$).

Discussion

Experiment 2 replicated the results of Experiment 1 using an indirect memory task. As hypothesized by the RMR conjecture, no recency-in-memory effect for the last memory encountered was observed when using melodies in an unfamiliar tuning system. This suggests that preventing recency-in-memory effects in memory for melody using an unfamiliar tuning system is not limited to conscious recognition task instructions, but instead generalizes to indirect memory task instructions. Consistent with Experiment 1 and not supporting the prediction of the RMR conjecture, the number of intervening items had no statistically significant disruptive effect. However, the non-significant effect could still be due to the small number of intervening melodies (up to 13).

It is also possible that listeners were capable of perceiving the stimuli as slight aberrations of melodies in the musical tradition they are familiar with. The stimuli used in Experiments 1 and 2 were based on European folk song melodies and therefore closely follow Western musical tradition on the basis of many musical features (e.g., contour, rhythm, equal temperament tuning). The main difference in the stimuli here, when compared with the original melodies in Herff et al. (submitted), is the step size of the tuning system. It appears this manipulation was sufficient to disrupt the recency-in-memory effect, but potentially the stimuli were too close to the familiar tuning system to disrupt entirely the integration of the melodies as a whole. Experiment 3 therefore investigates cumulative disruptive effects over a larger range of intervening melodies, using stimuli that are distinctively different to those that are typically familiar to Western listeners.

Experiment 3: disruptive memory effects at larger numbers of intervening items

Experiment 3 investigates the possibility that melodies in a distinctively unfamiliar tuning system elicit cumulative disruptive effects over large numbers of intervening melodies (beyond the 13 used in Experiments 1 and 2). The data from Experiment 3 are from a larger study conducted at Murdoch University, Australia. The study investigated statistical learning of artificial grammars in the context of music, specifically grammars applied to pitch and rhythm. During the procedure, participants were first exposed to melodies in an unfamiliar tuning system. The exposure phase in this project consisted of a continuous recognition paradigm. This is the same paradigm that Experiment 1, Experiment 2 and previous studies (e.g., Herff et al., submitted) used in the context of melody recognition. After the exposure phase, participants completed various follow-up experiments (data not reported here). The present investigation analyzed only the recognition data of the exposure phase.

During the continuous recognition paradigm of the exposure phase, participants listened to melodies in an unfamiliar tuning system presented three times throughout the experiment and provided recognition responses to each melody. The large number of intervening melodies (up to 107) and the two repetitions of each melody allow us to assess potential cumulative effects from the number of intervening items between both first and second, as well as second and third, presentations of a melody.

Method

Participants. In total, 105 participants (largely undergraduate students) were recruited from the Murdoch University community ($M_{age} = 25$ years, $SD_{age} = 7.5$, 36 males/68 females). Participants' musical training ranged from 0 to 17 years ($M_{training} = 2.1$ years, $SD_{training} = 3.3$, 32 participants with more than 2 years of musical training). Participation was either voluntary or in exchange of course credit.

Stimuli and equipment. The unfamiliar tuning system used here consisted of the following pitch heights: 480, 520, 560, 605 and 665 Hz. Note durations of 60, 110, 550 and 920 ms, with a 100-ms silent gap between notes, were used. Melodies consisted of five or six notes. All notes were synthesized pure tones with 10 ms linear onset and offset ramps. This resulted in melodies that did not conform to Western music tradition in rhythm and tonality.[2]

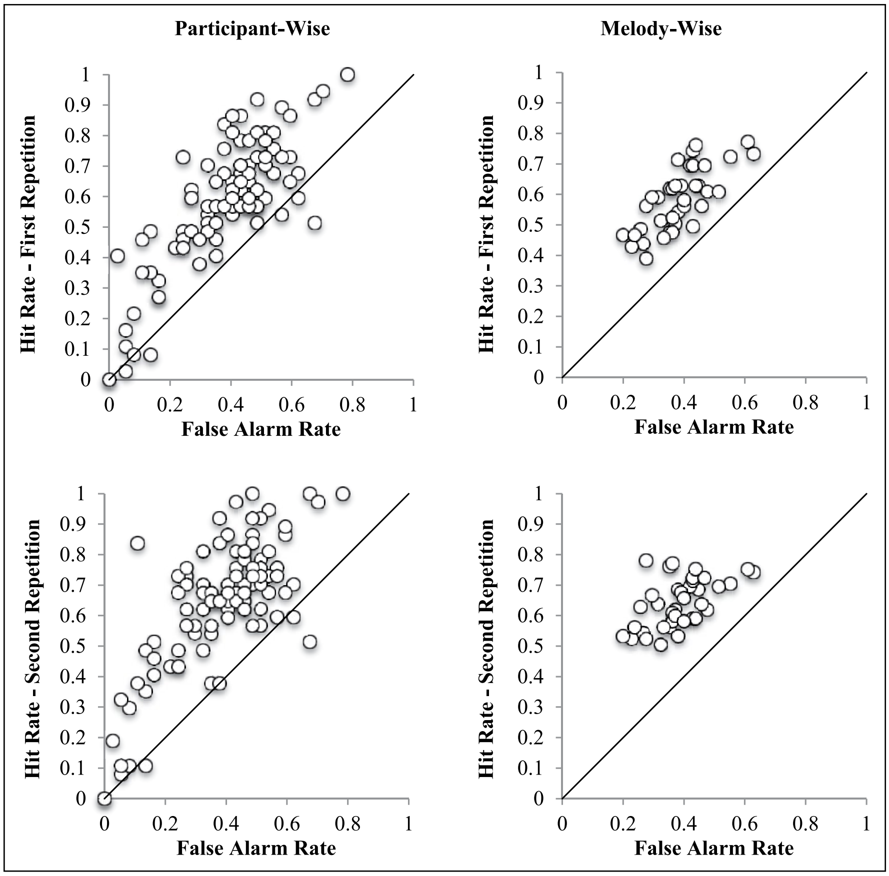

Figure 5. Hit rates and false alarm rates in Experiment 3. The left panels show the data participant-wise and the right panels melody-wise. The top row shows data from the first and second presentations (i.e., first repetition). The bottom row shows data from the first and third presentations (i.e., second repetition). The reference line represents chance level. Overall, the performance was significantly above chance and higher on the second repetition of the melodies.

Durations and pitch discriminability were piloted to ensure that all pitch and duration differences were clearly discriminable ($N=9$).[3] Using the tonal affinity model (Milne, 2013; Milne et al., 2011), this tuning system shows cosine similarity of 0.16186 to 12-TET and 0.22268 to the 88.08-CET system used in Experiments 1 and 2. This means the unequally tempered tuning system used in Experiment 3 is less similar to 12-TET than the tuning system used in Experiment 2.

Procedure. Participants gave informed consent, were seated in front of a computer and completed a short demographic questionnaire. Similar to Experiment 1, participants were instructed that they would hear many different melodies, one after another, and it was their task to indicate whether they had heard a given melody in this experiment before. Participants responded by pressing one of two keys on the keyboard. The keys were counterbalanced between participants. The presentation order of the melodies was fully randomized. Thirty-seven melodies appeared three times throughout the experiment for a total of 111 trials.

Results

Figure 5 shows melody- and participant-wise performance. Overall, participants performed significantly above chance ($Z=24.27$, $p<0.0001$).[4]

A generalized mixed-effects model investigated the hypothesis that, in contrast to Western-tonal melodies, melodies in an unfamiliar tuning system elicit cumulative disruptive effects of the number of intervening melodies. This hypothesis was supported between the first and second occurrences of a melody. A model predicting "old" responses on melody repetitions using a random intercept for *Subject* and *Melody*, as well as a systematic factor for *Dynamic Response Tendency* ($BIC=4777.2$, $LogLik=-2372.1$), showed a significant improvement when provided with the *Number of Intervening Items* ($BIC=4759.4$, $LogLik=-2359$, $p<0.0001$, $\Delta BIC=17.8$).

The hypothesis was also supported between the second and third occurrences of a melody. A model predicting "old" responses on melody repetitions using a random intercept for *Subject* and *Melody*, as well as a systematic factor for *Dynamic Response Tendency* ($BIC=4542.6$,

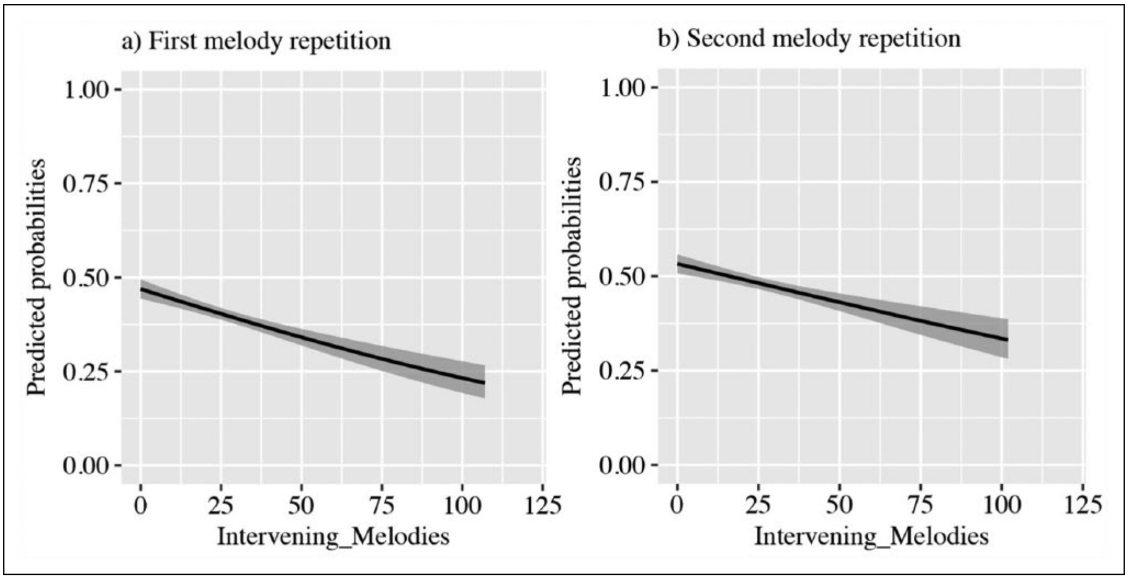

Figure 6. Prediction lines of generalized mixed-effects models that predict the probability of bias-corrected recognition (*y*-axis) in Experiment 3. The left panel shows the effect of the number of intervening melodies between the first and second presentations of the melodies. The right panel shows the effect between the second and third presentations. A significant disruptive effect of the number of intervening items on bias-corrected recognition performance is observed in both periods. The gray area around the prediction line represents a 95% confidence interval.

LogLik=−2254.8), significantly improved when provided with the *Number of Intervening Items* (*BIC*=4532.1, *LogLik*=−2245.4, $p<0.0001$, ΔBIC=10.5). Figure 6 depicts the bias-free probability of recognition responses as the number of intervening items increased between the first and second, as well as second and third, presentations of a melody.

Delay, number of intervening melodies or stimuli. Above, we have shown cumulative disruptive interference in Experiment 3 but not in Experiments 1 and 2. However, it is not clear whether the effect in Experiment 3 was driven by the larger number of intervening melodies, the larger temporal delay between first and second presentations of the melodies or the greater unfamiliarity of the musical system behind the stimuli. To disentangle these possible explanations, we conducted a secondary analysis of the 1-13 intervening melody conditions in Experiment 3 (to be directly comparable with Experiments 1 and 2). Observing cumulative disruptive interference in these conditions would show that the effect is due to the stimuli, rather than the larger temporal delay or a greater number of intervening items. This is because the stimuli in Experiment 3 are shorter compared to those used in Experiments 1 and 2 (thus providing the opportunity to exclude greater temporal delay as an explanation). Furthermore, this analysis used the same number of intervening melodies (thus providing the opportunity to exclude greater number of intervening melodies as an explanation).

Between the first and second presentations of the melodies, a model predicting “old” responses on melody repetitions using a random intercept for *Subject* and *Melody*, as well as a systematic factor for *Dynamic Response Tendency* (*BIC*=1588.6, *LogLik*=−779.94), showed a significant improvement when provided with the *Number of Intervening Items* (*Coef*: −0.05, $p=0.0013$; *BIC*=1585.4, *LogLik*=−774.77, $p=0.0013$, ΔBIC=3.2).

The same was observed between the second and third presentations of the melodies, where a model predicting “old” responses on melody repetitions using a random intercept for *Subject* and *Melody*, as well as a systematic factor for *Dynamic Response Tendency* (*BIC*=1625.5, *LogLik*=−798.4), showed a significant improvement when provided with the *Number of Intervening Items* (*Coef*: −0.01, $p<0.0001$; *BIC*=1617.2, *LogLik*=−790.68, $p<0.0001$, ΔBIC=8.3).

Discussion

Experiment 3 aimed to investigate cumulative effects at relatively large numbers of intervening melodies using an unfamiliar tuning system. It was hypothesized that with up to 107 intervening melodies tested here, a cumulative disruptive effect of intervening melodies should be observed. Such a disruptive effect is common in recognition in general (Bui et al., 2014; Donaldson & Murdock, 1968; Herff et al., submitted; Hockley, 1992; Konkle et al., 2010; Olson, 1969; Poon & Fozard, 1980; Rakover & Cahlon, 2001; Sadeh et al., 2014) but has not previously been observed in the context of melody recognition when a familiar tuning system with up to 195 intervening melodies is presented (Herff et al., submitted). Experiment 3 supported this hypothesis. Participants could perform the task, but a significant disruptive effect on recognition performance unfolded as the number of intervening melodies increased. This effect was still present when participants' shifts in response tendencies and random effects of melody were accounted for. A secondary analysis of the data suggests that the cumulative disruptive effects in Experiment 3 are predominantly driven by the stimuli rather than larger temporal delay or a greater number of intervening items.

General discussion

In general, a disruptive effect on memory from the number of intervening items between first and second presentations of a target stimulus has been observed in stimuli from a variety of domains (Bui et al., 2014; Donaldson & Murdock, 1968; Hockley, 1992; Konkle et al., 2010; Olson, 1969; Poon & Fozard, 1980; Rakover & Cahlon, 2001; Sadeh et al., 2014). However, recognition of melodies in a familiar tuning system does not appear to be affected by commonly reported interference from intervening items (Herff et al., submitted). In three experiments, this study further investigated this phenomenon by manipulating the number of intervening items between first and second presentations, as well as between second and third presentations of target melodies, sounded in an unfamiliar tuning system. A recent RMR conjecture (Herff et al., submitted) predicts a disruptive effect of the number of intervening items when melodies are presented in an unfamiliar, rather than familiar, tuning system. Furthermore, the conjecture predicts no recency-in-memory advantage for immediate melody repetition. The results here showed no recency-in-memory effect for recognition of melodies presented in unfamiliar tuning systems. A disruptive effect from the number of intervening melodies was only observed in Experiment 3 (up to 107 intervening melodies) and not in Experiments 1 and 2 (up to 13 intervening melodies). The present findings will now be discussed in the context of recency-in-memory and effects from the number of intervening melodies.

Effects from recency-in-memory

Recency-in-memory phenomena describe memory advantages for the last encountered item (Jahnke, 1963; Roberts, 1986). In the context of music, such an advantage for the last item has been previously demonstrated in recognition of single notes as well as melodies in a familiar tuning system (Berz, 1995; D. Deutsch, 1970, 1975; Dowling, 1973; Greene & Samuel, 1986; Herff et al., submitted). To explain recognition advantages for the last encountered melody with recency-in-memory effects, it is necessary to assume that melodies are perceived and represented as integrated entities (see D. Deutsch, 1986; Krumhansl, 1991). The RMR conjecture predicts that if a melody is not perceived and integrated as a whole, then only a recency-in-memory effect for the last underlying part of a melody remains (e.g., single notes or phrases). This was tested here using melodies in an unfamiliar tuning system designed to disrupt perception and integration of melodies as a whole.

Specifically, the RMR conjecture assumes that prior experience influences perception and that perception directly influences formation of new memories. If prior experience does not exist (e.g., of pitch structure in an unfamiliar tuning system) and therefore does not inform the process of integrating a melodic cluster of notes into a perceptually coherent melody, then no memory representation of a perceptually integrated melody will be formed. Instead, memory is based only on representations of the melody's underlying components. As a result of the missing representation of an integrated melody, a recency-in-memory effect for the last encountered melody should not emerge. The present data support this hypothesis. Recognition performance in Experiment 1 was statistically similar between immediate melody repetition and up to 3 intervening melodies. In Experiment 2, the change in perceived familiarity was also statistically similar. These results also support the finding of previous studies that show recency-in-memory effects and effects from the number of intervening items on memory for melody are similarly captured in explicit recognition tasks and indirect perceived familiarity tasks (Herff et al., submitted).

Of relevance to future investigations is an unpublished pilot study conducted by the first author that tested the design of Experiment 1 in a small sample ($N = 14$) of musicians (> 2 years of formal training), in contrast to the less experienced participants in Experiment 1. Interestingly, the sample of musicians also did not show a recency-in-memory effect. Empirical testing of populations with varying degrees of musical expertise or exposure to specific tuning systems is an important avenue for further research. The RMR conjecture predicts that regardless of tuning system, listeners familiar with the system should show a recency-in-memory effect, whereas listeners with no experience in the system should show no recency-in-memory advantage. This could provide a phenomenon-driven approach to explore cultural differences in the perception of auditory stimuli.

Effects from the number of intervening melodies

As predicted by the RMR conjecture, the hypothesized disruptive effect from the number of intervening items was observed in Experiment 3. This experiment used up to 107 intervening melodies. In contrast, Experiments 1 and 2 investigated intervening gap sizes of only up to 13 melodies, and neither produced significant disruptive effects from the number of intervening items. Interestingly, in Experiment 3 disruptive effects from the number of intervening items were observed between the first and second and between the second and third presentations of the melodies. These results show that, once observed, the effect is fairly robust even when repetition strengthens representations.

The methodological differences between Experiment 3 and the previous experiments provide alternative explanations for our findings. First, Experiment 3 used large numbers of intervening items, whereas Experiments 1 and 2 investigated a relatively small number. Second, Experiment 3 used a different melody corpus. The melodies in Experiments 1 and 2 were played with a piano timbre and used pitch contours and rhythms common to melodies in familiar music traditions such as Western-tonal music. In Experiment 3, the melodies were played using pure tones rather than a musical timbre. However, this difference in timbre most likely did not drive the effects because, so far, only differences in vocal timbre have been shown to impact memory when compared to other musical timbres (Weiss, Vanzella, Schellenberg, & Trehub, 2015). Nevertheless, the effect of timbre on memory for melodies is a topic that deserves greater attention in future studies.

The unfamiliar tuning system utilized in Experiments 1 and 2 was the most dissimilar equal-tempered tuning system compared to 12-TET within the given range. However, it is still an equal-tempered tuning system and therefore carries some perceptual similarity to 12-TET. Indeed, the cosine similarity measures reveal that the non-equal tempered tuning system used in Experiment 3 is decidedly more dissimilar to 12-TET than the tuning system used in Experiments 1 and 2. Experiment 3 also used a far less musical timbre (pure tones), and rhythms were not matched directly to common and familiar rhythms in Western-tonal music as they were in Experiments 1 and 2. This means that they did not adhere to regular metrical structures and overall pitch contour. The stimuli in Experiment 3 can therefore be defined as far more extreme in terms of unfamiliarity than the stimuli in Experiments 1 and 2. It could be that the tuning system in Experiments 1 and 2 was unfamiliar enough to disruptive recency-in-memory effects, but not unfamiliar enough to induce effects from the number of intervening melodies (at least when the number was $\leq$13 melodies). The melodies in Experiment 3, being far less familiar than those in Experiments 1 and 2, were likely unfamiliar enough to disrupt integration entirely.

To further investigate which factors may explain the effects reported in Experiment 3, we performed a secondary analysis of the 1-13 intervening melody conditions in Experiment 3. We found cumulative disruptive interference in these conditions, even though the stimuli in Experiment 3 were shorter than those in Experiments 1 and 2. This result shows that the cumulative disruptive effect in Experiment 3 is indeed, as discussed above, stimulus-driven rather than by the larger number of intervening melodies and the associated greater temporal delay.

These findings suggest that the issue of listeners' unfamiliarity with different musical features and its effect on recency-in-memory and disruptive effects from intervening items is a fruitful future research endeavor. In general, the RMR conjecture predicts smaller effects from recency-in-memory and more of a disruptive effect from the number of intervening items as a tuning system's unfamiliarity increases. A future systematic manipulation of individual musical features while measuring the aforementioned memory phenomena has the potential to shed light on fundamental memory processing in general and how previous experience influences perception and formation of subsequent memories in particular. This line of investigation has begun by investigating the relative influence of pitch information and rhythm information on memory for melodies (Herff et al., in press) and will continue by systematically manipulating individual musical features and their associated impact on memory.

Conclusion

The aim of this study was to deepen our understanding of memory for melody by investigating the RMR conjecture using recognition of melodies in an unfamiliar tuning system. Previously, recognition of melodies in a familiar tuning system has shown recency-in-memory effects, but not cumulative disruptive effects from the number of intervening melodies (Herff et al., submitted). The RMR conjecture predicts opposite findings for melodies in an unfamiliar tuning system. The present data support these predictions using explicit as well as indirect recognition tasks: no recency-in-memory effects were observed, but cumulative disruptive effects from the number of intervening melodies in memory for melodies in an unfamiliar tuning system were observed only when the number of intervening items increased from 13 up to 107.

The findings of this study help advance our understanding of memory's stimulus specificity (Fougnie, Zughni, Godwin, & Marois, 2015) by exploring why some domains show clear disruptive effects from the number of intervening items, whereas others do not. However, future research is needed to investigate precise degrees of prior experience and its effect on both recency-in-memory and interference from the number of intervening items. Much research is also needed to extend the present findings to stimuli beyond the domain of music. Finally, the RMR conjecture appears to be a useful conceptual tool to guide future research investigating the critical link between prior experience, perception and subsequent formation of memory representations. An investigation of how prior experience influences perception, how this perception determines formation of representations and how these representations are then affected by basic memory phenomena has the potential to shed important light on the fundamental relationship between human memory, perception and, ultimately, day-to-day behavior.

Acknowledgements

We thank Andrew Milne and the MARCS Institute Music Cognition and Action group for valuable comments on a previous version. We thank Esra Mungan for the constructive suggestion to specifically analyze the first 13 intervening melody conditions in Experiment 3 in order to pinpoint what drives our findings. We thank Barbara Tillmann, Mari Jones and Kate Stevens for their involvement in the artificial grammar study that, as a side product, produced the data analyzed here, as well as their willingness to share data. We thank Donna Caporn, Rebekah Hunt, Christopher Innes, Farah Jaffar, Harry Malcom, Prachi Singh and Sarah Tan-Nygard who assisted in data collection.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

Supplemental material

The Supplemental Material is available at journals.sagepub.com/doi/suppl/10.1177/1747021817734978

Notes

1. Note that the musical features have been calculated on the original versions of the melodies.
2. The melodies were constructed in various artificial grammars, unfamiliar to listeners, which were of importance for the later investigation of statistical learning that will be reported elsewhere.
3. In pitch perception and a standard/comparison task with a silent retention interval, a Weber fraction of 0.04 led to discrimination performance above 90%. The pitch Weber fractions used in the present stimuli were between 0.08 and 0.10, therefore clearly discriminable. In duration perception, Weber fractions of 0.30 led to discrimination performance above 98%. The duration fractions used in this study ranged from 0.67 to 1.27, therefore also clearly discriminable.
4. Note that the here-observed cumulative disruptive interference between the first and second, as well as second and third, presentations of the melodies is also observed in both groups when participants with fewer or more than 2 years of musical training are analyzed separately (all $p < 0.05$).

References

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. New York, NY: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. doi:10.1016/J.Jml.2007.12.005

Barascud, N., Pearce, M. T., Griffiths, T. D., Friston, K. J., & Chait, M. (2016). Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *Proceedings of the National Academy of Sciences*, *113*, E616–E625.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. doi:10.1016/j.jml.2012.11.001

Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2013). Lme4: Linear mixed-effects models using Eigen and s4 (*R package version*, 1, 2-23). Retrieved from https://cran.rproject.org/web/packages/lme4/index.html.

Berch, D. B. (1976). Criterion change in continuous recognition memory: A sequential effect. *Bulletin of the Psychonomic Society*, *7*, 309–312. doi:10.3758/BF03337199

Berman, S., Friedman, D., & Cramer, M. (1991). ERPs during continuous recognition memory for words and pictures. *Bulletin of the Psychonomic Society*, *29*, 113–116.

Berz, W. L. (1995). Working memory in music: A theoretical model. *Music Perception*, *12*, 353–364.

Buchsbaum, B. R., Padmanabhan, A., & Berman, K. F. (2011). The neural substrates of recognition memory for verbal information: Spanning the divide between short- and long-term memory. *Journal of Cognitive Neuroscience*, *23*, 978–991. doi:10.1162/Jocn.2010.21496

Bui, D. C., Maddox, G. B., Zou, F., & Hale, S. S. (2014). Examining the lag effect under incidental encoding: Contributions of semantic priming and reminding. *The Quarterly Journal of Experimental Psychology*, *67*, 2134–2148. doi:10.1080/17470218.2014.909506

Campeanu, S., Craik, F. I. M., Backer, K. C., & Alain, C. (2014). Voice reinstatement modulates neural indices of continuous word recognition. *Neuropsychologia*, *62*, 233–244. doi:10.1016/J.Neuropsychologia.2014.07.022

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of north India. *Journal of Experimental Psychology: General*, *113*, 394–412.

Cui, A. X., Collett, M. J., Troje, N. F., & Cuddy, L. L. (2015). Familiarity and preference for pitch probability profiles. *Cognitive Processing*, *16*, 211–218. doi:10.1007/s10339-015-0651-7

Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science*, *168*, 1604–1605. doi:10.1126/science.168.3939.1604

Deutsch, D. (1975). The organization of short-term memory for a single acoustic attribute. In D. Deutsch & J. A. Deutsch (Eds.), *Short-term memory* (pp. 107–151). New York, NY: Academic Press.

Deutsch, D. (1986). Auditory pattern recognition. In K. R. Boff, L. Kaufman & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. II: Cognitive Processes and Performance, pp. 32.1–32.49). New York, NY: Wiley.

Deutsch, D., & Boulanger, R. C. (1984). Octave equivalence and the immediate recall of pitch sequences. *Music Perception: An Interdisciplinary Journal*, *2*, 40–51.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review*, *70*, 80–90. doi:10.1037/h0039515

Donaldson, W., & Murdock, B. B. (1968). Criterion change in continuous recognition memory. *Journal of Experimental Psychology*, *76*, 325–330. doi:10.1037/H0025510

Dowling, W. J. (1973). Rhythmic groups and subjective chunks in memory for melodies. *Perception & Psychophysics*, *14*, 37–40.

Dowling, W. J. (1991). Tonal strength and melody recognition after long and short delays. *Perception & Psychophysics*, *50*, 305–313. doi:10.3758/Bf03212222

Fougnie, D., Zughni, S., Godwin, D., & Marois, R. (2015). Working memory storage is intrinsically domain specific. *Journal of Experimental Psychology: General*, *144*, 30. doi:10.1037/a0038211

Friedman, D. (1990). Cognitive event-related potential components during continuous recognition memory for pictures. *Psychophysiology*, *27*, 136–148.

Gaudreau, D., & Peretz, I. (1999). Implicit and explicit memory for music in old and young adults. *Brain and Cognition*, *40*, 126–129.

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*, 517–524.

Halekoh, U., & Højsgaard, S. (2014a). A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models: The R package pbkrtest. *Journal of Statistical Software*, *59*, 1–32.

Halekoh, U., & Højsgaard, S. (2014b). Pbkrtest: Parametric bootstrap and kenward roger based methods for mixed model comparison (R package version 0.4-0). Retrieved from http://CRAN.R-project.org/package=pbkrtest

Halpern, A. R., & Bartlett, J. C. (2010). Memory for melodies. In M. R. Jones, R. R. Fay & A. N. Popper (Eds.), *Music perception* (Vol. 36, pp. 233–258). New York, NY: Springer. doi:10.1007/978-1-4419-6114-3_8

Halpern, A. R., & O'Connor, M. G. (2000). Implicit memory for music in Alzheimer's disease. *Neuropsychology*, *14*, 391–397. doi:10.1037/0894-4105.14.3.391

Herff, S. A., Olsen, K. N., & Dean, R. T. (2017). Resilient memories for melodies: The number of intervening melodies does not influence novel melody recognition. *Quarterly Journal of Experimental Psychology*. doi:10.1080/17470218.2017.1318932

Herff, S. A., Olsen, K. N., Prince, J., & Dean, R. T. (2017). Interference in memory for pitch-only and rhythm-only sequences. *Musicae Scientiae*. doi:10.1177/1029864917695654

Hockley, W. E. (1992). Item versus associative information: Further comparisons of forgetting rates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1321–1330. doi:10.1037/0278-7393.18.6.1321

Jahnke, J. C. (1963). Serial position effects in immediate serial recall. *Journal of Verbal Learning and Verbal Behavior*, *2*, 284–287. doi:10.1016/S0022-5371(63)80095-X

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*, 54–69. doi:10.1037/A0028347

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795. doi:10.1080/01621459.1995.10476572

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, *53*, 983–997. doi:10.2307/2533558

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General*, *139*, 558–578. doi:10.1037/a0019165

Krumhansl, C. L. (1991). Music psychology: Tonal structures in perception and memory. *Annual Review of Psychology*, *42*, 277–303. doi:10.1146/Annurev.Ps.42.020191.001425

Krumhansl, C. L., & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 579.

Kruschke, J. K. (2010). Null hypothesis significance testing. In J. K. Kruschke (Ed.), *Doing Bayesian data analysis: A tutorial introduction with R* (pp. 215–238). New York, NY: Academic Press.

Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science*, *6*, 299–312. doi:10.1177/1745691611406925

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, *142*, 573–603. doi:10.1037/a0029146

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception*, *22*, 663–714.

Milne, A. J. (2013). *A computational model of the cognition of tonality*. Milton Keynes, UK: The Open University.

Milne, A. J., & Holland, S. (2016). Empirically testing Tonnetz, voice-leading, and spectral models of perceived triadic distance. *Journal of Mathematics and Music*, *10*, 59–85.

Milne, A. J., Laney, R., & Sharp, D. B. (2015). A spectral pitch class model of the probe tone data and scalic tonality. *Music Perception*, *32*, 364–393. doi:10.1525/Mp.2015.32.4.364

Milne, A. J., Laney, R., & Sharp, D. B. (2016). Testing a spectral model of tonal affinity with microtonal melodies and inharmonic spectra. *Musicae Scientiae*, *20*, 465–494.

Milne, A. J., Sethares, W. A., Laney, R., & Sharp, D. B. (2011). Modelling the similarity of pitch collections with expectation tensors. *Journal of Mathematics and Music*, *5*, 1–20.

Nathoo, F. S., & Masson, M. E. J. (2016). Bayesian alternatives to null-hypothesis significance testing for repeated-measures designs. *Journal of Mathematical Psychology*, *72*, 144–157.

Norman, D. A. (2013). *Models of human memory*. New York, NY: Elsevier.

Oberauer, K., & Lewandowsky, S. (2011). Modeling working memory: A computational implementation of the time-based resource-sharing theory. *Psychonomic Bulletin & Review*, *18*, 10–45. doi:10.3758/S13423-010-0020-6

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review*, *19*, 779–819. doi:10.3758/S13423-012-0272-4

Olson, G. M. (1969). Learning and retention in a continuous recognition task. *Journal of Experimental Psychology*, *81*, 381–384. doi:10.1037/h0027756

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, *76*, 241–263. doi:10.1037/h0027272

Parncutt, R., & Cohen, A. J. (1995). Identification of microtonal melodies: Effects of scale-step size, serial order, and training. *Perception & Psychophysics*, *57*, 835–846. doi:0.3758/Bf03206799

Pearce, M. T. (2014). IDyOM project. Retrieved from https://code.soundsoftware.ac.uk/projects/idyom-project

Poon, L. W., & Fozard, J. L. (1980). Age and word-frequency effects in continuous recognition memory. *Journals of Gerontology*, *35*, 77–86. doi:10.1093/geronj/35.1.77

R Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available from http://www.R-project.org/

Rakover, S. S., & Cahlon, B. (2001). *Face recognition: Cognitive and computational processes* (Vol. 31). Amsterdam, The Netherlands: John Benjamins.

Roberts, L. A. (1986). Modality and suffix effects in memory for melodic and harmonic musical materials. *Cognitive Psychology*, *18*, 123–157. doi:10.1016/0010-0285(86)90010-1

Sadeh, T., Ozubko, J. D., Winocur, G., & Moscovitch, M. (2014). How we forget may depend on how we remember. *Trends in Cognitive Sciences*, *18*, 26–36. doi:10.1016/J.Tics.2013.10.008

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52. doi:10.1016/S0010-0277(98)00075-4

Schellenberg, E. G. (1996). Expectancy in melody: Tests of the implication-realization model. *Cognition*, *58*, 75–125.

Schellenberg, E. G., & Habashi, P. (2015). Remembering the melody and timbre, forgetting the key and tempo. *Memory and Cognition*, *43*, 1021–1031. doi:10.3758/s13421-015-0519-1

Schneider, A. (1997). “Verschmelzung,” tonal fusion, and consonance: Carl Stumpf revisited. In M. Leman (Ed.), *Music, gestalt, and computing: Studies in cognitive and systematic musicology* (pp. 117–143). New York, NY: Springer.

Schon, D., & Francois, C. (2011). Musical expertise and statistical learning of musical and linguistic structures. *Frontiers in Psychology*, *2*, 167. doi:10.3389/fpsyg.2011.00167

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*, 461–464.

Shepard, R. N., & Teghtsoonian, M. (1961). Retention of information under conditions approaching a steady-state. *Journal of Experimental Psychology*, *62*, 302–309. doi:10.1037/H0048606

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*, 34–50. doi:10.1037/0096-3445.117.1.34

Stroop, J. R. (1992). Studies of interference in serial verbal reactions (reprinted from journal experimental-psychology, vol 18, pp. 643-662, 1935). *Journal of Experimental Psychology: General*, *121*, 15–23.

Tillmann, B., & McAdams, S. (2004). Implicit learning of musical timbre sequences: Statistical regularities confronted with acoustical (dis)similarities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1131–1142. doi:10.1037/0278-7393.30.5.1131

Weiss, M. W., Vanzella, P., Schellenberg, E. G., & Trehub, S. E. (2015). Pianists exhibit enhanced memory for vocal melodies but not piano melodies. *The Quarterly Journal of Experimental Psychology*, *68*, 866–877.

Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics*, *9*, 60–62. doi:10.2307/2957648

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*, 547–558. doi:10.1038/nrn2152