New Ideas in Psychology 55 (2019) 68–77

Contents lists available at ScienceDirect

New Ideas in Psychology

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

Investigating cumulative disruptive interference in memory for melodies, words, and pictures

Steffen A. Herff[a,b,*], Kirk N. Olsen[c,d], Aydin Anic[c,d], Nora K. Schaal[e]

[a] The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Australia
[b] Institute of High Performance Computing, Agency for Science, Technology and Research, Singapore
[c] Department of Psychology, Macquarie University, Australia
[d] Centre for Elite Performance, Expertise, and Training, Macquarie University, Australia
[e] Department of Experimental Psychology, Heinrich-Heine University Düsseldorf, Germany

ARTICLE INFO

Keywords:
Memory
Interference
Music perception
Recognition
Regenerative Multiple Representations

ABSTRACT

Cumulative disruptive interference in memory describes the well-established phenomenon that recognition performance decreases as the number of intervening items between first and second stimulus presentation increases. Memory for melody has been shown to exhibit resilience to this type of recognition interference, and a novel Regenerative Multiple Representation (RMR) conjecture has been developed to explain these findings. Here, we critically assess, replicate, and extend key findings and predictions of the RMR conjecture. In four tasks ($N = 68$), we critically test whether memory for melodies' (Task 1) resilience to cumulative disruptive interference holds when compared to memory for pictures (Task 2) and words (Task 3–4) when many of the previous analytical and methodological discrepancies within the literature are accounted for. Furthermore, we test a prediction of the RMR conjecture that words written in a plain sans-serif style of writing (Task 3) should show stronger cumulative disruptive interference compared to words in an elaborate longhand style of writing (Task 4). Lastly, we explored potential auditory context effects of noisy unintelligible multi-talker babbling on memory for melodies, pictures, and words. As hypothesized, we found strong cumulative disruptive interference in recognition for written words and pictures, but not for melodies. The predicted differences between the two styles of writing was not supported. However, we found evidence for a domain-dependent auditory context effect that can be explained by an increase in cumulative disruptive interference in mix-matching contexts when encoding occurred under adverse listening conditions, but retrieval did not. The findings provide support for some of the assumptions and predictions of the RMR conjecture, and pave the way for future studies that utilise the RMR conjecture as a theoretical framework for understanding the intimate relationship between memory and perception.

1. Introduction

An important aspect of memory that may not receive the attention it deserves is the question of how we forget, rather than how we remember. Two main contributors to forgetting are *Decay* and *Interference* (Eysenck & Keane, 2015*)*. Decay refers to a decrease in memory performance simply due to the passing of time. Interference, on the other hand, refers to reduced memory performance due to cumulative disruptive effects of additional information that affect a memory trace. For example, accurately recognizing the driver of a taxi becomes substantially more difficult if you tend to take many taxis every day. Cumulative disruptive interference is commonly observed in a wide range of stimuli across different modalities (Sadeh, Ozubko, Winocur, & Moscovitch, 2014). Indeed, seemingly ubiquitous recognition of digits, word lists, or faces tends to decrease as the number of intervening items increases between the first instance of a target item and when it reappears (Buchsbaum, Padmanabhan, & Berman, 2011; Bui, Maddox, Zou, & Hale, 2014; Campeanu, Craik, Backer, & Alain, 2014; Deutsch, 1970, 1975; Donaldson & Murdock, 1968; Friedman, 1990b; Hockley, 1992; Konkle, Brady, Alvarez, & Oliva, 2010; Olson, 1969; Poon & Fozard, 1980; Rakover & Cahlon, 2001; Sadeh et al., 2014). Considering that cumulative disruptive interference is widely observed across multiple domains, every exception that does not show this type of interference deserves special attention because they help shed light on mechanisms that underpin effective long-term recognition. Recently, Herff, Olsen, and Dean (2018) demonstrated that memory for melodies

---

[*] Corresponding author.1 Fusionopolis Way, #16Connexis North, 138632, Singapore|
E-mail address: s.herff@westernsydney.edu.au (S.A. Herff).

https://doi.org/10.1016/j.newideapsych.2019.04.004
Received 29 December 2018; Received in revised form 17 April 2019; Accepted 22 April 2019

is one such example that shows little to no susceptibility to cumulative disruptive interference. For example, brief target melodies (~12s) presented throughout a listening experiment are recognized equally well regardless of whether one or 194 intervening melodies are presented between the first and second presentation of a target. A novel Regenerative-Multiple-Representations (RMR) conjecture was developed by Herff and colleagues to explain these findings (detailed below). Considering the novelty of the conjecture, many findings require replication, extension, and critical assessment. Therefore, the present work was designed to critically assess the assumptions of the RMR conjecture by replicating and extending previous findings supporting the conjecture. Specifically, we aim to: (1) replicate memory for melodies' resilience towards cumulative disruptive interference; (2) critically assess in a series of tasks whether memory for melodies shows resilience from cumulative disruptive interference when compared with other non-musical stimuli such as pictures or written words; (3) replicate previous findings of an auditory context effect on cumulative disruptive interference (i.e., background noise vs. quiet); and (4) extend these findings by exploring whether this effect extends to cross-modal situations. In the following, we will elaborate on the RMR conjecture and the aims and rationale for the series of tasks utilised in the present study.

1.1. Regenerative-Multiple-Representations

The RMR conjecture describes a crucial link between prior experience, perception, and formation of new memories. It assumes that prior experience influences the way in which sensory input is perceived and interpreted. In turn, it is the perception and interpretation that is subjected to memory (Herff, Olsen, & Dean, 2018). The conjecture draws from several key experimental findings and established memory theories (e.g., Paivio, 1969). One key finding that inspired this conjecture is that words and photographs show cumulative disruptive interference in memory (for words, see Bäuml, 1996; McGeoch, 1932; Underwood, 1957; for photographs, see Deffenbacher, Carr, & Leu, 1981; Konkle et al., 2010; Nickerson, 1965), whereas poetry and simple line object drawings, similar to the aforementioned findings with melodies, do not (Berman, Friedman, & Cramer, 1991; Friedman, 1990a; Tillmann & Dowling, 2007). An intriguing commonality between the stimulus domains that do not show cumulative disruptive effects are strong interdependent connections between underlying components that are integrated into a coherent whole; a process previously termed *perceptual synthesis* (Deutsch, 1986, 2013). Indeed, the RMR conjecture predicts that music, poetry, and drawings are perceived as a set of underlying components that are integrated into a coherent whole. For example, a melody consists of underlying components such as notes, intervals, and rhythms. Given familiarity with the underlying rules (i.e., the statistical occurrence) of how melodies' underlying components inter-relate, these components can be integrated into a coherent whole: a melody. Importantly, in the case of melodies, both the underlying components and the coherent whole are remembered. It is argued that this is a crucial difference between music, poetry, and drawings – all of which show resilience towards cumulative disruptive effects from intervening items – and their counterparts of spoken word, prose, and photographs, all of which do show cumulative disruptive effects in memory. It appears that in stimuli such as spoken word, prose, and photographs, a memory representation is formed from the integrated whole, whereas the underlying components are rapidly forgotten (Babcock & Freyd, 1988; Deutsch, 1986; Knoblich, Seigerschmidt, Flach, & Prinz, 2002; Krumhansl, 1991; Rayner, Pollatsek, Ashby, & Clifton Jr, 2012a, 2012b; Schneider, 1997; Tse & Cavanagh, 2000). So why might we remember only the coherent whole for some stimuli, but remember a coherent whole *as well as* the underlying components for other stimuli?

The RMR conjecture posits that the reason lies in prior experience. Over the course of a lifetime, observers learn the most relevant way of perceiving a stimulus (Goldstone, 1998). A similar consideration has previously been described as *pertinence*, a weighting of perception based on a current situation and a long-term factor such as prior knowledge (Deutsch, 1986, see also J. A. Deutsch & Deutsch, 1963). Indeed, experimental studies show that prior knowledge influences and guides perception (Bruner & Postman, 1949; Bülthoff, Bülthoff, & Sinha, 1998; Goldstone, 1995; Malmberg & Annis, 2012). Returning to why this is relevant for the lack of cumulative disruptive effects in music, poetry, and drawings, we suggest that these are all stimuli in which humans have learned to pay close attention to the integrated whole and its underlying components. For example, in prose, the semantic information is by far the most important information; underlying words are of less relevance (Rayner et al., 2012b). In poetry, on the other hand, the precise underlying wording is similarly important to the overarching semantic and affective meaning.

Given that prior knowledge informs multiple perception of stimuli like music, poetry, and drawings, multiple memory representations are formed. Importantly, these representations have redundancy as they partially code the same information. Similar to other theories of memory (see dual-coding theory by Paivio, 1969), the RMR conjecture predicts that multiple representations assist in retrieval by regeneration of lost information. The candidate mechanism in the context of memory for melodies are the strong expectancies observed in music that guide listeners' attention, help in predicting what is coming next, and may be used to interpolate forgotten parts of the music (Margulis, 2005; Pearce, 2014; Schellenberg, 1996). For example, we perceive the underlying components of a melody and because we are familiar with the 'rules', we form an integrated, coherent representation of the melody as a whole (Cui, Collett, Troje, & Cuddy, 2015; Saffran, Johnson, Aslin, & Newport, 1999; Schon & Francois, 2011; Tillmann & McAdams, 2004). However, if we are not familiar with these rules, the RMR conjecture predicts that we cannot form a representation of the stimulus as an integrated whole.

Furthermore, the RMR conjecture postulates that it is multiple representations of a melody that provides memory for melody with a resilience from cumulative interference observed in Herff, Olsen, and Dean (2018). This means that if the formation of integrated representations of a simple melody is disrupted, then the previously observed resilience from cumulative interference should disappear. One way of disrupting a listeners' representation of a melody as a whole is by presenting melodies in a tuning system unfamiliar to the listener. Unfamiliar tuning systems provide a context where the listener is unlikely to be familiar with the 'rules' of how underlying components of such melodies inter-relate. This prediction of the RMR conjecture was tested in Herff, Olsen, Dean, and Prince (2018), and as predicted, melodies that previously showed no cumulative disruptive interference when played in a familiar tuning system *did show* cumulative disruptive interference when played in an unfamiliar tuning system. In other words, it is likely that melodies in an unfamiliar tuning system are only perceived as their underlying components (i.e., a series of disconnected notes) because the necessary experience required to integrate the components into a coherent melody is lacking. On the other hand, melodies in a familiar tuning system are perceived as an integration of their underlying components (a coherent melody) as well as the underlying components (such as individual notes, intervals, and phrases). Thus, it is likely that melodies in a familiar tuning system develop resilience against cumulative interference in memory due to the additional representations listeners can form through their prior experience with the tuning system.

Moreover, if melodies in an unfamiliar tuning system show cumulative disruptive interference when they consist of a combination of different underlying components (e.g., a dynamic pitch-sequence coupled with a dynamic rhythm-sequence), then the RMR conjecture predicts that the underlying components themselves should show cumulative disruptive interference when tested separately in melodies that consist of only a dynamic pitch-sequence or only a dynamic rhythm-sequence. This is because the separate underlying components provide fewer possible percepts and memory representations than a melody that combines multiple sets of underlying components. This hypothesis was also tested and supported in Herff, Olsen, Dean, et al. (2018). Taken together, these findings have provided preliminary support for the RMR conjecture and have begun to shed greater light on whether memory for music is 'special', and why (Stevens, 2015).

1.2. *Extension of the RMR conjecture beyond music*

Currently, the RMR conjecture is a useful tool to inspire further research, especially in, but not limited to, the domain of music perception. It is capable of making clear, falsifiable predictions in a variety of domains that can formally test the generalisability of its key tenets. For example, the RMR conjecture assumes clear cumulative disruptive interference should be observed in pictures. This is because observers perceive the underlying components of a picture and integrate them into a coherent whole (i.e., the subject of the picture). In this case, a memory representation of the integrated whole is formed. In the case of pictures - as opposed to music - the underlying components are generally speaking of little relevance. As a result, fewer representations are formed, rendering memory for pictures susceptible to cumulative disruptive interference. This is tested in Task 2 in the present study (Memory for Pictures). Similarly, written words should show cumulative disruptive interference. This is because readers tend to rapidly integrate the underlying components (e.g., letters) into a coherent whole (a semantic word) but do not pay much attention to the individual letters (Rayner et al., 2012a). Indeed, it is a common finding that letters are skipped during reading in favour of a rapid integrated semantic whole (or word) (Rayner et al., 2012b).

Furthermore, words written in longhand should provide less cumulative disruptive interference when compared to words written in plain type font without ornaments (e.g., serifs) (Herff, Olsen, & Dean, 2018). This prediction derives from the fact that readers track the direction as well as the mode of production of words written in longhand. Therefore, elaborate handwriting (similar to drawings) provides the opportunity for representations that goes beyond the integrated concept (the meaning of the word) (Babcock & Freyd, 1988; Knoblich et al., 2002; Tse & Cavanagh, 2000). This prediction is tested here in Task 3 (plain type font) and 4 (ornamented long hand). If there is stronger cumulative disruptive interference in Task 3 relative to Task 4, then this would provide further evidence for the RMR conjecture.

An important issue always remained about the RMR conjecture: to compare the lack of cumulative disruptive interference in memory for melodies with the strong cumulative disruptive interference using other stimuli, results using musical stimuli were compared to previously published literature that used non-musical stimuli but also different memory paradigms. With this comes the question of comparability. For example, most studies have used a block-type design study, rather than the continuous recognition paradigm used in the studies that developed the RMR conjecture. Furthermore, other studies tested far greater, or smaller numbers of stimuli, or participants (Konkle et al., 2010; Poon & Fozard, 1980). Some previous studies also tested cumulative disruptive interference at fixed intervals rather that allowing the number of intervening items to vary freely across the task. Last but not least, every research group may use slightly different ways of assessing statistical relevance of a given effect. Here, we will test in a controlled environment cumulative disruptive interference in melodies (Task 1), pictures (Task 2) and words (Tasks 3–4).

1.3. *Context-effects in cumulative interference in memory*

As the present study tests cumulative disruptive interference with stimuli from musical and non-musical domains, it presents an opportunity to expand on another recent related finding. As discussed above, there is no cumulative disruptive interference in memory for melody when the melodies are learned and tested with the same auditory background. However, cumulative disruptive interference does occur when the auditory context between learning and testing changes (Herff, Dean, & Schaal, 2018). Herff et al. presented listeners with 'noisy' unintelligible multi-talker babbling concurrently with the presentation of a melody during either a melody's first presentation but not the second (Noise-first presentation/Clear-second presentation), its second presentation but not the first (Clear-first presentation/Noise-second presentation), during both first and second presentation (Noise-first presentation/Noise-second presentation), or during neither presentation (Clear-first presentation/Clear-second presentation). In this design, the background noise between learning and test trials was either congruent (Clear-Clear and Noise-Noise) or incongruent (Clear-Noise and Noise-Clear). Overall, recognition performance was worse when background noise was present. Importantly, cumulative disruptive interference was descriptively stronger for incongruent conditions (Clear-Noise and Noise-Clear) and interference only reached significance when noise was presented during melodies' first presentation but not during their second presentation (Noise-Clear). Indeed, not even the intuitively most disruptive condition (Noise-Noise) showed significant cumulative disruptive interference. Within the framework of the RMR conjecture, stronger interference in mismatching contexts seems plausible. This is because the background noise during melodies' first presentation disrupts later recognition, however, it also provides an additional percept for listeners that can help form a memory representation. If the second presentation of a melody also contains background noise (Noise-Noise), then the additional representation may help in regenerating the target melody, thus providing some resilience against cumulative disruptive interference when compared to a condition without background noise during the melodies' second presentation (Noise-Clear). The present study further tests these hypotheses in the context of memory for melodies, pictures, and words.

1.4. *Overview of the study*

This study aimed to further test the central tenets of the RMR conjecture in four tasks by: (a) replicating and extending the research by Herff, Olsen, and Dean (2018) on memory for melodies using pictures and words; and (b) investigating whether memory for non-auditory stimuli such as pictures and words is negatively affected by background noise, similar to that of the melodies reported in Herff, Dean et al. (2018). It was hypothesized that: (1) memory for melodies will show no cumulative interference (Task 1); (2) memory for pictures will show cumulative interference (Task 2) as will memory for words (Task 3 and 4); and (3) words written in longhand (Task 4) will show less cumulative interference in memory than words written in plain-type font (Task 3). The question of whether the auditory context (background noise) affects cumulative interference in memory across modalities was addressed throughout all four tasks. This issue was more exploratory in nature, hence no directional hypotheses.

2. Method

2.1. *Participants*

Participants were recruited from Macquarie University in Australia ($N = 44$, $M_{age} = 21.4$, $SD_{age} = 6.5$, female = 27, male = 17, other = 0) as well as English proficient students from the Heinrich-Heine-University in Germany ($N = 24$, $M_{age} = 21.3$, $SD_{age} = 10.9$, female = 21, male = 3, other = 0).[1] The combined sample of 68 participants ($M_{age} = 22.1$, $SD_{age} = 5.9$, female = 48, male = 20, other = 0) had an average of 2.9 years of musical training ($SD = 4.1$). All 68

---

[1] The present study was not concerned with questions regarding cross-cultural comparisons, thus data from the German and the Australian sample were aggregated and analysed together.

participants completed the melody task (Task 1) and the picture task (Task 2). Of the total 68 participants, 36 participants (M_{age} = 21.3, SD_{age} = 4.6, female = 26, male = 10, other = 0) were randomly allocated to perform the word recognition task with the plain type font and 32 participants (M_{age} = 23.0, SD_{age} = 6.9, female = 22, male = 10, other = 0) performed the task with the elaborate type font. The order in which participants completed the tasks was quasi-randomized to ensure that all possible order combinations were equally distributed across the study.

2.2. Stimuli and equipment

Task 1 used 40 randomly chosen melodies from the corpus previously used in Herff, Dean, and Olsen (2017) and Herff, Olsen, and Dean (2018). The exact same 40 melodies used here were also previously used in Herff, Dean, et al. (2018). The melodies (all ~12 s in duration) were presented through headphones and participants could adjust the volume to their preferred level. Online supplemental material S1 contains all stimuli in .wav and midi format, as well as a musical feature analysis of each melody.

Task 2 presented 40 pictures of waves previously used in Konkle et al. (2010) from the category ‘scene-wave’ obtained through http://cvcl.mit.edu/MM/sceneCategories.html. The exact same 40 pictures were also previously used in (Wessendorf, 2018).

Task 3 and Task 4 both presented 40-word stimuli. These words were taken from the first forty items of the combined A and B list of the Rey Auditory Verbal Learning Test (RAVLT) (Schmidt, 1996). The words were presented on screen clearly visible in font size 45 for 5 s. Task 3 presented the words in the ‘Arial’ font. Arial is a widely used sans-serif font that represent the plain computer font in the present experiment. Task 4 presented words in the ‘Austen’ font. ‘Austen’, created by Pia Frauss, is an elaborate font rich with ornaments designed to represent Jane Austen's handwriting found in personal letters (Frauss, 2005). The font can be obtained from https://www.dafont.com/jane-austen.font and was used here to represent a longhand font in contrast to the Arial sans-serif font. Examples of both fonts can be seen in Fig. 1.

The unintelligible multi-talker background noise manipulation was incorporated in all four tasks and was identical to the stimuli used in Herff, Dean, et al. (2018). Specifically, a number of short snippets were taken from a continuous 15-min sound file of multi-talker noise, examples can be found in the online supplement file S1 (Davis & Kim, 2012; Davis, Kim, Grauwinkel, & Mixdorff, 2006; Munhall, Jones, Callan, Kuratate, & Vatikiotis-Bateson, 2004). No snippet was used twice in the study to ensure no repetition in the background multi-talker noise. The MAX/MSP software platform (Version 7.3.5) was used to present the experiments' protocol, response tools and audio playback.

2.3. Continuous recognition paradigm

All four tasks used a continuous recognition paradigm (Shepard & Teghtsoonian, 1961). In this paradigm, participants are continuously presented with stimuli, one at a time. Each stimulus presentation is considered a trial and after the presentation of each stimulus, participants are required to respond whether they think they have encountered that particular stimulus before in the present experiment. The next trial is initiated as soon as the participant provided their response. In each task 40 stimuli were presented. Every stimulus was presented twice throughout the tasks, resulting in a total of 80 trials per task. The number of intervening items until a target item reappeared was fully randomized. This means that every participant in every task received a unique randomization list that was automatically created on each experiment start-up. A schematic representation of the paradigm used here is shown in Fig. 2.

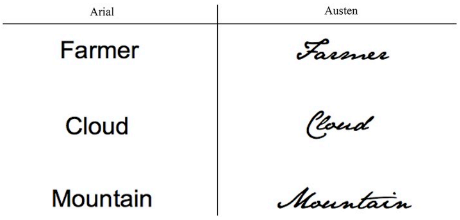

Fig. 1. This figure depicts a subset of the words used in the experiment. The left column shows words in Arial font; the right column shows words using the Jane Austen font.

In each task, concurrent unintelligible multi-talker babbling was also presented either during the first presentation of a stimulus but not the second (Noise-first presentation/Clear-second presentation), its second presentation but not the first (Clear-first presentation/Noise-second presentation), during both first and second presentation (Noise-first presentation/Noise-second presentation), or during neither presentation (Clear-first presentation/Clear-second presentation) (Herff, Dean, et al., 2018). The four different background multi-talker noise conditions are shown in Table 1.

2.4. Procedure

After providing informed consent, each participant completed a demographic questionnaire and then began three of the four tasks, the order of which were quasi-randomized and determined prior to commencement. For each task, participants were presented with one melody (Task 1), picture (Task 2), or word (Task 3 and 4) per trial, and were asked whether they had heard or seen that stimulus before in the task. They achieved this by clicking on either an ‘old’ or ‘new’ button on the computer screen after each stimulus presentation. The instructions were amended to reflect the specific kind of stimulus presented in each task. Each melody was 12 s in duration in Task 1, followed by a prompt for participants to provide their 'Old/New' response.[2] In Task 2, each picture was presented on the computer monitor for 5 s before disappearing. This procedure was equivalent for word presentation in Task 3 and 4. After the third and final task, participants were debriefed on the purpose of the study. Overall the study took approximately 45 min to complete.

2.5. General statistical approach

The statistical approach closely follows previous work with continuous recognition paradigms (Herff & Czernochowski, 2019; Herff, Olsen, & Dean, 2018; Herff, Olsen, Prince, & Dean, 2018). We use generalised linear mixed effects models across the four tasks to analyse data (Baayen, 2008; Baayen, Davidson, & Bates, 2008; Judd, Westfall, & Kenny, 2012; Kass & Raftery, 1995; Kruschke, 2010, 2013; Nathoo & Masson, 2016). In the results section for each task, we first established whether participants performed significantly above chance by reporting the coefficient estimate, z-value, and p-value of an *Occurrence* factor that codes first vs. second melody presentation in a generalised mixed effects model that takes random participant response biases and melody variation into account. If the factor predicts significantly more “old” response during melodies' second presentation, then participants overall where capable of performing the memory task.

To explore cumulative disruptive interference as well as context effects, we used models that predict binary coded recognition response (‘Old’ or ‘New’) of responses during second stimulus presentation. Responses during first stimulus presentation were used to capture participants' response biases and bias shifts as detailed later. The

---

[2] After each response, participant also rated their confidence on a 100-point vertical visual analog scale. Data reported elsewhere.

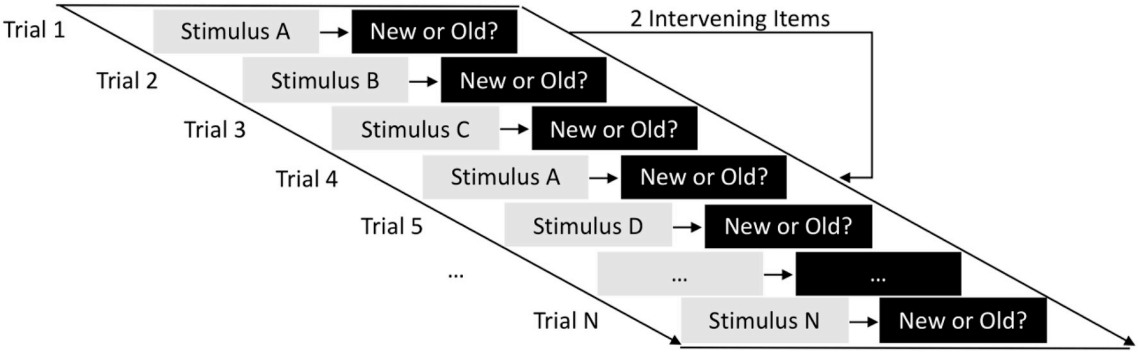

Fig. 2. Schematic of the continuous recognition paradigm deployed here. On each trial, participants are presented with a stimulus and are asked whether or not they have heard this stimulus in this task before. In the example above, Stimulus A is first presented in Trial 1 and presented a second time in Trial 4. In Trial 1, the correct answer would be 'New' since it is Stimulus A's first presentation. In Trial 4, after two intervening items, the correct answer would be 'Old', since this constitutes Stimulus A's second presentation.

Table 1
The four multi-talker background noise conditions used in all four tasks.

		Stimulus Second Presentation	
		Clear	Noise
Stimulus First Presentation	Clear	Clear-Clear	Clear-Noise
	Noise	Noise-Clear	Noise-Noise

Note. The term 'Noise' indicates that presence of multi-talker background noise presented during the first or second presentation of a specific melody (Task 1), picture (Task 2) or word (Task 3 and Task 4). The term 'Clear' indicates no presence of multi-talker background noise. The order of 'Noise' and 'Clear' in the table refers to the presence or absence of noise during the first presentation of a stimulus (the first word in a word-pair), or the presence or absence of noise during the second presentation of a stimulus (the second word in a word-pair).

models consisted of random intercepts for *Participant* and *Melody* (Barr, Levy, Scheepers, & Tily, 2013). A fixed factor for *Number of Intervening Items* implements cumulative disruptive interference between first and second presentation of an item. A fixed factor for *Noise Condition* (Clear-Clear, Clear-Noise, Noise-Clear, Noise-Noise) implements potential context effects of background noise. The models were built in the R environment (R-Core-Team, 2013) using the lme4 package (Bates, Maechler, Bolker, & Walker, 2013).

To further assess the models, we compared models using likelihood-ratio tests (Wilks, 1938). This step prevents the capture of significant effects that are only due to an increase in model complexity (Kruschke, 2011). We report delta Bayes Information Criteria (BIC) (Schwarz, 1978) by subtracting the BIC of the model with the additional factor, from a model without the factor. We consider a ΔBIC of two or greater as "positive" evidence in favour of the model with the additional *Number of Intervening Items* factor. A ΔBIC difference of six or greater is considered as "strong" evidence (Kass & Raftery, 1995).

An important - yet often overlooked - concern when deploying a continuous recognition paradigm are dynamic shifts in participant response tendencies over the course of the task (Berch, 1976; Donaldson & Murdock, 1968; Snodgrass & Corwin, 1988). Similar to previous work (Herff, Olsen, & Dean, 2018; Herff, Olsen, Dean, et al., 2018), we account for such shifts, by training participant-wise generalised mixed effects models on 'Old' responses on first presentations based on trial number. The models were then used to predict the probability of pressing 'Old' on second presentation trials. The resulting predictions were implemented as the fixed *Dynamic Response* in all models analysing memory performance. Effectively, these models control for a shift in the tendency to provide False Alarms over the course of the task. Below, we will provide detail about the individual tasks where they differ from the general methodology.

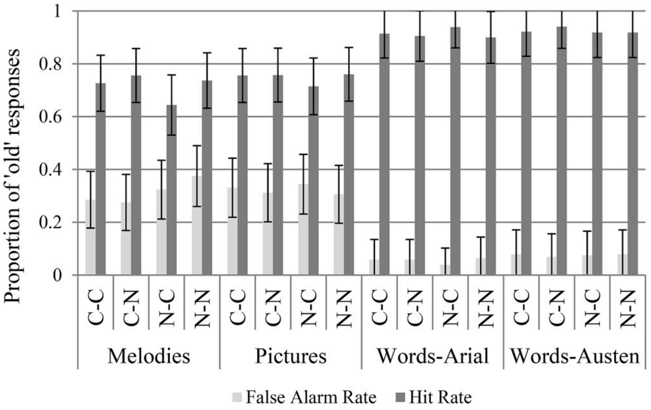

Fig. 3. False Alarm and Hit Rates for all background noise conditions and tasks. False Alarm Rates (light grey) were lower than Hit Rates (dark grey) across all tasks and conditions showing that participants were able to perform the recognition tasks. Both word as well as the picture recognition task showed no differences between the four background noise condition; Clean-Clean (C–C), Clean-Noise (C–N), Noise-Clean (N–C), and Noise-Noise (N–N). The melody task, however, showed significantly reduced performance when background noise was present during encoding, but not retrieval. This can be visualized by the reduced distance between False Alarm Rate and Hit Rate in the Melody task in the Noise-Clean (N–C) condition. Error bars indicate 95% CIs.

3. Results

3.1. Task 1: melody recognition

Overall, participants performed significantly above chance ($Est = 1.904, z = 29.321, p < .0001$). This can also be seen in Fig. 3 by the significantly higher proportion of overall 'old' responses during first presentation (False Alarm Rate) compared to second presentation (Hit Rate). Fig. 3 shows Hit Rates and False Alarm rates for all tasks, split by background noise condition.

The Number of Intervening Items did not disrupt performance ($\text{Est} = -0.005$, $\text{SE} = 0.003$, $z = 1.877$, $p = .065$, $\Delta\text{BIC} = -4.4$). Note that the negative ΔBIC (i.e., greater ΔBIC in the model with the additional Number of Intervening Items fixed factor) also indicates that a model without the Number of Intervening Item outperforms a model with the additional information. This supports the notion that information about the number of intervening items does not help in predicting melody recognition performance.

Model comparison revealed that Noise Context significantly increased model performance ($\Delta\text{BIC} = 3$, $p < .001$). Specifically, this effect was due to reduced performance in the Noise-Clear condition ($\text{Est} = -0.456$, $\text{SE} = 0.128$, $z = 3.560$, $p < .001$). Direct comparison

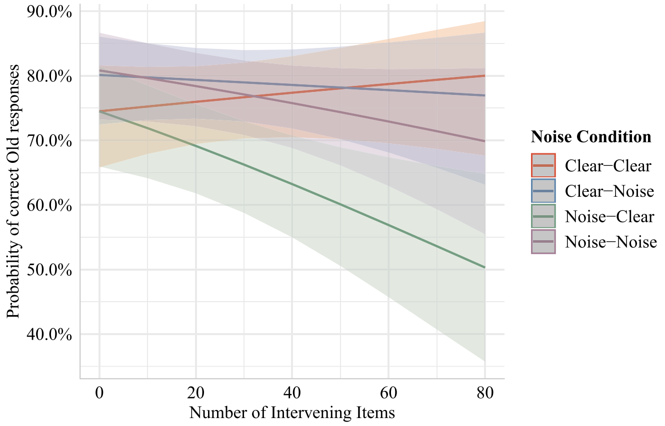

Fig. 4. Marginal effect plots of cumulative disruptive interference in the melody recognition task across background noise conditions. The Clear-Clear, Clear-Noise, and Noise-Noise condition show no significant cumulative disruptive interference. The Noise-Clear condition, however, shows significant cumulative disruptive interference. As the number of intervening items increases, the probability of producing a correct 'Old' response decreases in the Noise-Clear condition. The transparent areas around the prediction lines represent 95% Confidence bands.

showed that performance in Noise-Clear was significantly lower compared to Clear-Clear (Est = 0.456, SE = 0.128, z = 3.559, p < .0001), Clear-Noise (Est = 0.613, SE = 0.130, z = 4718, p < .0001), as well as Noise-Noise (Est = 0.515, SE = 0.129, z = 4.001, p < .0001). The other three conditions, Clear-Clear, Clear-Noise, and Noise-Noise did not differ significantly from one another (all p > .226). Further investigation of the Number of Intervening Items x Noise Condition Interaction showed that the reduced performance in the Noise-Clear condition was due to an increase in cumulative disruptive interference (Est = −0.017, SE = 0.007, z = −0.464, p = .014). This can be seen in Fig. 4.

As hypothesised, no significant cumulative disruptive interference was observed in the melody recognition task. Furthermore, as predicted, we found a significant effect of background context. The Noise-Clear condition led to significantly lower performance compared to the Clear-Clear condition. This performance decrement appeared to be driven by exacerbated cumulative disrupted interference in the Noise-Clear condition that was not present during the other background noise conditions.

3.2. Task 2: picture recognition

As shown in Fig. 3, overall, participants performed significantly above chance (Est = 2.067, z = 30.849, p < .0001). The Number of Intervening Items significantly disrupted performance (Est = −0.039, SE = 0.003, z = −12.455, p < .0001, ΔBIC = 159.6). Model comparison revealed that information about Noise Context did not significantly increase model performance (ΔBIC = −18.5, p = .155). This is also visible in Fig. 3. The negative ΔBIC lends additional support that information about the Noise-Condition does not help in predicting picture recognition performance. As hypothesised, participants showed strong cumulative disruptive interference in the picture recognition task. Furthermore, the background-noise did not affect picture recognition performance.

3.3. Task 3: word recognition (arial-font)

As shown in Fig. 3, overall, participants performed significantly above chance (Est = 5.473, z = 30.893, p < .0001). The Number of Intervening Items significantly disrupted performance (Est = −0.035, SE = 0.005, z = −6.403, p < .0001, ΔBIC = 37.17). Model comparison revealed that information about Noise Context did not significantly increase model performance (ΔBIC = −17.06, p = .190). The negative ΔBIC lends additional support that information about the Noise-Condition does not help in predicting word recognition performance. This can also be seen in Fig. 3.

Similar to the picture recognition task, but dissimilar to the melody recognition task, the word recognition task showed strong cumulative disruptive interference. Again, similar to the picture recognition task, the background-noise context did not affect memory performance. Importantly, Task 3 used a plain sans-serif font and it was hypothesized that cumulative disruptive interference would be stronger in the plain sans-serif font compared to a more elaborate longhand used in Task 4.

3.4. Task 4: word recognition (austen-font)

As shown in Fig. 3, overall, participants performed significantly above chance (Est = 5.996, z = 27.985, p < .0001). The Number of Intervening Items significantly disrupted performance (Est = −0.036, SE = 0.006, z = −6.002, p < .0001, ΔBIC = 29.98). Model comparison revealed that information about Noise Context did not significantly increased model performance (ΔBIC = −19.87, p = .661). The negative ΔBIC lends additional support that information about the Noise-Condition does not help in predicting word recognition performance. This is also visible in Fig. 3. When analysed together with the data from Task 3, the interaction term Number of Intervening Items x Font did not reach significance (Est = 0.001, SE = 0.008, z = 0.13, p = .896).

The results of Task 4 mirror those from Tasks 2 and 3. A strong cumulative disruptive interference is observed as the number of intervening items increases. Simultaneously, the background noise context exhibits no detectable effects on word recognition performance. Previously the RMR conjecture predicted that words written in elaborate longhand should be more resilient to cumulative disruptive interference than words written in plain sans-serif type font. We did not find support for this prediction.

3.5. Cumulative disruptive interference across tasks

A generalised mixed effects model combining the data of all tasks specifically investigates differences between the tasks in terms of their cumulative disruptive interference with a Number of Intervening Items × Task interaction term. The picture recognition task (Task 2: Est = −0.025, SE = 0.004, z = −7.214, p < .0001), and the Word recognition task using the Arial font (Task 3: Est = −0.029, SE = 0.005, z = −5.299, p < .0001), as well as the Austen font (Task 4: Est = −0.024, SE = 0.006, z = −3.986, p < .0001) all showed significantly stronger cumulative disruptive interference than the melody recognition task (which did not show significant interference on its own). The cumulative disruptive interference between the two Word and the Picture recognition task did not significantly differ from one another (all p-values > .53). Cumulative disruptive interference for all four tasks is shown in Fig. 5.

4. Discussion

The RMR conjecture was proposed to explain the recent finding that some stimuli such as melodies do not show cumulative disruptive interference, whereas stimuli in other domains such as pictures and words do. Here, we critically tested several key assumptions and predictions of the RMR conjecture within and outside the domain of music. In four tasks testing memory for melodies, pictures, and words, we found that memory for melodies remains remarkably resilient to cumulative disruptive interference, whereas memory performance for words and pictures decreased as the number of intervening items increased. We also tested whether words written in plain type font show more cumulative disruptive interference than words written in elaborate longhand. This hypothesis was not supported. Lastly, we replicated previous

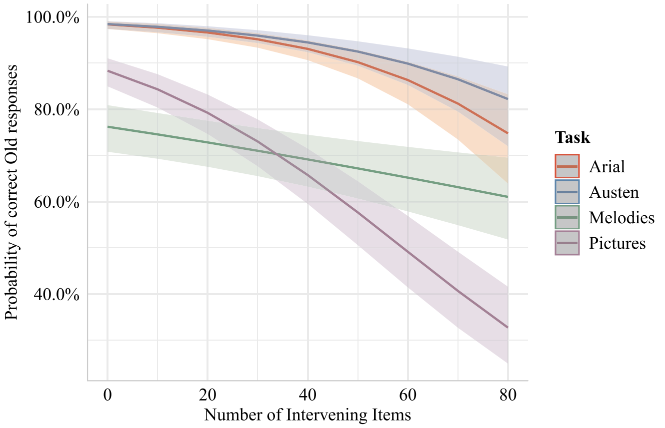

Fig. 5. Marginal plots of the cumulative disruptive interference in the different tasks. Overall, both word recognition tasks as well as the picture recognition task showed significant cumulative disruptive interference. The melody recognition task, however, did not show significant cumulative disruptive interference. The transparent areas around the prediction lines represent 95% Confidence bands.

findings of an auditory background-context effect in memory for melody and showed that it does not generalise to other non-auditory stimuli. In the following section, we will discuss the individual findings in greater detail and conclude with future directions regarding the RMR conjecture.

4.1. *Cumulative disruptive interference*

In all four tasks, participants performed significantly above chance showing that participants were capable of correctly identifying which stimuli had been presented before, and which ones had not. However substantial differences in cumulative disruptive interference were observed between the tasks. Both word recognition tasks (Tasks 3 and 4) as well as the picture recognition task (Task 2) showed strong disruption of memory performance as the number of intervening items increased. This cumulative disruptive interference is well documented across a plethora of stimulus domains (Bui et al., 2014; Campeanu et al., 2014; Deutsch, 1970, 1975; Donaldson & Murdock, 1968; Friedman, 1990b; Hockley, 1992; Konkle et al., 2010; Nickerson, 1965; Olson, 1969; Poon & Fozard, 1980; Rakover & Cahlon, 2001, see Sadeh et al., 2014, for a review) and confirms prior findings on cumulative disruptive interference for words and pictures in Poon and Fozard (1980) and Konkle et al. (2010) respectively. Memory for melody, on the other hand, showed no significant cumulative disruptive interference. This result replicates previous findings in memory for melodies (Herff, Olsen, & Dean, 2018). Importantly, cumulative disruptive interference in the two word recognition tasks and the picture recognition task were comparable and substantially stronger than in the melody recognition task. Such a cross-stimulus comparison was previously only conducted between studies, separated by differences in paradigm, methodology, and analysis. Here, we show that when controlling for many of the differences that previously made accurate cross-stimulus comparisons difficult, melodies continue to show resilience to cumulative disruptive interference whereas other stimuli do not. Interestingly, this is the case even when general memory performance was better in the word recognition tasks compared to the melody recognition task (see Fig. 3). We take the present findings as further support of the RMR conjecture.

4.2. *Type of word font*

Previously, the RMR conjecture predicted that words written in plain sans-serif type font (e.g., Arial font) should show stronger cumulative disruptive interference compared to words written in elaborate hand writing (Austen font). Similar to pencil drawings, words written in elaborate longhand provide a percept of the direction as well as mode of production (the underlying components), in addition to the integrated meaning of the word (Babcock & Freyd, 1988; Knoblich et al., 2002; Tse & Cavanagh, 2000). The additional percept of the underlying components should form additional memory representations that provide some resilience against cumulative disruptive interference. However, we did not find support for this prediction. Indeed, significant cumulative disruptive interference was observed in both the plain Arial and the elaborate Austen type font, yet interference between the two fonts was statistically indistinguishable. This finding was not predicted by the RMR conjecture and should be further investigated. Several possible explanations are pertinent here. First, Arial is one of the most commonly used fonts on digital as well as printed documents. Potentially a practice effect of Arial could have provided additional resilience to the cumulative disruptive interference, compensating for any initial difference between the two fonts. Furthermore, while the Austen font was based on actual handwriting and is undeniably more elaborate than the Arial font, it is still a digital font that consists of a small set of samples for each letter, rather than the variability encountered in free longhand on paper. Overall performance in memory for words in both font types was also approaching ceiling. The same number of word stimuli may create a lower cognitive load than picture or melody stimuli. Future research could consider digitalising actual handwriting in longhand, rather than using a digital font that approximates handwriting in longhand, while presenting a greater number of stimuli. In addition, a less commonly used sans-serif type font or set of nonsensical words may help elucidate differences in cumulative disruptive interference between font types.

4.3. *The role of auditory background*

Throughout the four tasks, half the trials were accompanied with unintelligible multi talker babbling in the background. The background noise was presented either in the first presentation of a stimulus during encoding (Noise-Clear), the second presentation of a stimulus during retrieval (Clear-Noise), both first and second presentation of a stimulus (Noise-Noise), or neither (Clear-Clear). In the melody recognition task, the Noise-Clear condition showed significantly lower performance compared to all other conditions. In previous studies, the Noise-Clear condition was also the only condition that significantly differed from the Clear-Clear condition, however, the other comparisons only approached significance (Herff, Dean, et al., 2018). In the present study, all comparisons between the Noise-Clear condition and the other conditions were significant. This may be due to greater power from the larger sample size used in the present study (68 here vs. 40 in Herff, Dean, et al., 2018).

That the Clear-Noise condition leads to lower performance compared to the Clear-Clear condition is intuitive, simply because background noise in general is associated with disruptive effects on memory (Gilbert, Chandrasekaran, & Smiljanic, 2014; Mattys, Davis, Bradlow, & Scott, 2012). However, background noise disrupted memory performance significantly more during encoding (Noise-Clear) compared to background noise during retrieval (Clear-Noise), showing an asymmetry of disruption depending on when the background noise is present during the memory process. This replicates the findings of Herff, Dean et al. (2018) and mirrors the findings from studies that use non-musical stimuli in divided attention paradigms (Anderson, Craik, & Naveh-Benjamin, 1998; Craik, Naveh-Benjamin, Ishaik, & Anderson, 2000; Fernandes & Moscovitch, 2000; Iidaka, Anderson, Kapur, Cabeza, & Craik, 2000; Naveh-Benjamin, Craik, Guez, & Dori, 1998; Naveh-Benjamin, Craik, Perretta, & Tonev, 2000; Naveh-Benjamin, Kilb, & Fisher, 2006; Park, Smith, Dudley, & Lafronza, 1989). A commonly provided explanation assumes that encoding and retrieval are two distinct processes, with encoding more prone to disruption than retrieval (Naveh-Benjamin et al., 2006).

The Noise-Clear/Noise-Noise comparison in the melody task is of particular interest. The finding that the Noise-Clear condition shows significantly lower performance compared to the Noise-Noise condition mirrors findings using spoken words rather than melodies (Creel, Aslin, & Tanenhaus, 2012), showing an auditory background context effect that enhances memory performance when study and test context are identical (Baddeley, Eysenck, & Anderson, 2009, pp. 176–180; Mattys, Bradlow, Davis, & Scott, 2013, pp. 75–82; Smith & Vela, 1992, 2001). Here, we provide further insight into this effect by tracing the reason of reduced performance in the Noise-Clear condition to an increase in cumulative disruptive interference. Indeed, when the background context was incongruent, and noise was presented during the first but not the second melody presentation, significant cumulative disruptive interference was observed. In all other conditions, there was no cumulative disruptive interference, leading to overall lower performance in the Noise-Clear condition. It is important to highlight that the Noise-Clear condition elicits lower performance compared to the Noise-Noise condition, whereas Clear-Noise condition does not elicit lower performance compared to the Clear-Noise condition. This suggests that it is not simply a universal context effect that is at play. In the context of the RMR conjecture, this finding implies that while additional background noise during encoding has a disruptive effect on recognition, it also offers formation of an additional percept (melody in noise). In the case where the second presentation of a stimulus was made under comparable background noise, the additional representation may assist in regenerating memory of the stimulus as a whole (i.e., the melody). This process ultimately results in better memory performance by providing additional resilience against interference (Herff, Dean, et al., 2018). However, as a clear encoding condition does not provide additional information that could be matched during retrieval, this effect becomes asymmetrical and only applies to disruption during encoding.

Across all tasks, the aforementioned context effects were only observed in the melody recognition task. Both word recognition tasks as well as the picture recognition task did not show a context effect induced by background babbling. In fact, no disruptive effect of background noise was observed in the first place. This supports the notion of a domain specific auditory background context effect, rather than a domain independent effect. The present results suggest that if auditory stimuli were encountered during a noisy background, then reinstating this background may facilitate memory performance. This carries implications for ear-witnesses who are tasked to recognize auditory events that happened in the past in noisy environments. Importantly, our data suggests that reinstating the auditory context may facilitate memory for auditory stimuli, but not for non-auditory stimuli. This could be further investigated by instantiating an inherently disruptive background (like the background babbling here) in the visual domain (e.g., visual blurring) and testing for enhanced cumulative disruptive interference in incongruent background conditions, specifically in the Noise-Clear condition.

4.4. Future directions for the RMR across modalities

The link between prior experience, perception, and formation of new memories provides a framework for future predictions regarding memory and perception tasks, independent of modality. For example, the RMR conjecture suggests that high domain-specific expertise is often associated with differences in perception, relative to low domain-specific expertise. For example, an expert carpenter perceives details of a chair that may be irrelevant for a normal perceiver; an expert architect can accurately identify and perceive combinations of architectural styles; an expert botanist is capable of identifying differences between plant species, perceptually invisible for uninformed observers; an expert entomologist can distinguish species that would be labelled identically by normal observers. Some of these changes in perception occur rapidly after being exposed to key information, whereas others require years of training. For example, Asian and African elephants can be hard to differentiate. However, if one learns that Asian elephants have two bumps on their heads whereas African elephants have one, the two species are easily distinguishable. As a result of newly acquired knowledge, perception changes in the future and future memory representations of elephants are enriched with the information of whether an elephant was African or Asian. This additional memory representation then aids memory formation. For example, if someone asks, ‘How many bumps were on the elephant that you saw earlier?‘, one's visual memory representation may have faded beyond the point where one can remember the exact number of bumps. However, the perceiver may still remember that they saw an Asian elephant, which can help ‘regenerate’ the information that the elephant had two bumps. In this example, a small piece of information changes perception and, subsequently, the formation of new memory representations.

Other changes in perception require intense training. Learning to perceive and produce pronunciation differences in an unfamiliar language, for example, can take a long time (Browman & Goldstein, 1995; Escudero & Chládková, 2010). In the framework of the RMR conjecture, the above examples are linked by the fact that additional experience (instructed or learned over long exposure) changes how an individual perceives the world. The RMR conjecture assumes that these additional percepts form additional memory representations that increase resilience to cumulative disruptive effects. Therefore, domain specific expertise can increase memory performance; however, it is important to note that the increased resilience toward intervening items is independent of overall memory performance. Future research could test this assumption by comparing experts and non-experts in one domain, predicting that experts show less cumulative disruptive interference. Furthermore, since prior information changes perception, it will be useful in future to design a paradigm that provides participants with additional information capable of fundamentally changing perception (e.g., syntactic rules in an artificial language). The way in which the additional information affects recognition performance and interference could provide insight into the crucial link between prior experience, perception, and memory formation.

Another useful future implication of the RMR conjecture is that it provides a framework to investigate ‘perceptual relevance’ of different stimulus features. Whether or not a specific feature of a stimulus has any implication for the observer is often difficult to establish. This is because perceptual relevance can only be defined based on the task (e.g., perceptually relevant to memory performance). The tasks to choose to gain insight into the general perceptual relevance of a stimulus is non-trivial. The link between prior experience, perception, and memory formation described by the RMR suggests that a minimum of two tasks is required. For example, one task that tests a measurable implication of a stimulus feature on memory, and one task that tests a measurable implication of a stimulus feature on evaluation. This is because it is the process of stimulus perception that forms a memory. If a specific stimulus feature shows a systematic influence on memory (either or both memory performance and memory bias), the feature by definition must have been perceived to cause this effect. This would be a clear sign of perceptual relevance. On the other hand, perception is informed by prior experience. Prior experience not only influences how a stimulus is perceived, but also provides a reference to how a stimulus is evaluated after being perceived. This means that if a stimulus feature shows a systematic effect on evaluation (e.g., liking), the specific feature must have interacted with an observer's prior experience. Perceptual relevance of this feature would thereby be established. This means a memory and an evaluation task can establish perceptual relevance of a stimulus feature by exploring the feature's interactions with memory formation and prior experience. Importantly, it is possible for a feature to systematically impact memory but not evaluation, and vice versa. Therefore, a memory and an evaluation task compensate for the shortcomings of the respective other in isolation, together providing a useful framework to investigate perceptual relevance of a stimulus feature.

5. Conclusion

The present work investigated an important part of memory – forgetting – across multiple modalities. Specifically, in four tasks we critically assessed underlying assumptions of the RMR conjecture. We found robust support for the resilience of memory for melodies against cumulative disruptive interference, whereas other stimuli such as words and pictures showed decreasing memory performance with each additional intervening item. In the word recognition task, the hypothesized difference in interference between a plain Arial font and a more elaborate handwritten Austen font was not supported and requires further empirical investigation. Lastly, we replicated an auditory context effect on cumulative disruptive interference in memory: whilst disruptive background noise during encoding impedes recognition performance, matching the same background noise during retrieval can enhance recognition. This effect was explained by an increase in cumulative disruptive interference when background noise was present during stimulus encoding, but not retrieval. Furthermore, the present data suggest that this effect does not generalise across modalities and thus is domain specific. Overall, these findings provide some support for the assumptions and predictions of the RMR conjecture, and pave the way for future studies that assess the utility of the RMR conjecture as a theoretical framework for understanding important memory processes.

Acknowledgements

This research was supported by a Macquarie University Centre for Elite Performance, Expertise, and Training (CEPET) Seed Grant awarded to SAH, KNO, and AA. Parts of the Discussion derive from an unpublished doctoral dissertation written by SAH. We would like to thank Lauren Fairley for constructive feedback on an earlier version of this manuscript.

References

Anderson, N. D., Craik, F. I. M., & Naveh-Benjamin, M. (1998). The attentional demands of encoding and retrieval in younger and older adults: I. Evidence from divided attention costs. *Psychology and Aging, 13*(3), 405.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* New York: 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*(4), 390–412. https://doi.org/10.1016/J.Jml.2007.12.005.

Babcock, M. K., & Freyd, J. J. (1988). Perception of dynamic information in static handwritten forms. *American Journal of Psychology,* 111–130.

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2009). *Memory.* New York: Psychology Press.

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*(3), 255–278. https://doi.org/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 4.

Bäuml, K. H. (1996). Revisiting an old issue: Retroactive interference as a function of the degree of original and interpolated learning. *Psychonomic Bulletin & Review, 3*(3), 380–384. https://doi.org/10.1037/0278-7393.7.4.299 and Memory, 7(4), 299-305.

Berch, D. B. (1976). Criterion change in continuous recognition memory: A sequential effect. *Bulletin of the Psychonomic Society, 7*(3), 309–312. https://doi.org/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*(2), 113–116.

Browman, C. P., & Goldstein, L. (1995). *Dynamics and articulatory phonology. Mind as motion.* 175–193.

Bruner, J. S., & Postman, L. (1949). Perception, cognition, and behavior. *Journal of Personality, 18*(1), 14–31.

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*(4), 978–991. https://doi.org/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. *Quarterly Journal of Experimental Psychology, 67*(11), 2134–2148. https://doi.org/10.1080/17470218.2014.909506.

Bülthoff, I., Bülthoff, H., & Sinha, P. (1998). Top-down influences on stereoscopic depth-perception. *Nature Neuroscience, 1*(3), 254.

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. https://doi.org/10.1016/J.Neuropsychologia.2014.07.022.

Craik, F. I. M., Naveh-Benjamin, M., Ishaik, G., & Anderson, N. D. (2000). Divided attention during encoding and retrieval: Differential control effects? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(6), 1744.

Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2012). Word learning under adverse listening conditions: Context-specific recognition. *Language & Cognitive Processes, 27*(7–8), 1021–1038.

Cui, A. X., Collett, M. J., Troje, N. F., & Cuddy, L. L. (2015). Familiarity and preference for pitch probability profiles. *Cognitive Processing, 16*(2), 211–218. https://doi.org/10.1007/s10339-015-0651-7.

Davis, C., & Kim, J. (2012). Is speech produced in noise more distinct and/or consistent? *Speech Sci. Technol.* 46–49.

Davis, C., Kim, J., Grauwinkel, K., & Mixdorff, H. (2006). Lombard speech: Auditory (A), visual (V) and AV effects. *Paper presented at the proceedings of the third international conference on speech prosody*.

Deffenbacher, K. A., Carr, T. H., & Leu, J. R. (1981). Memory for words, pictures, and faces - retroactive interference, forgetting, and reminiscence. *J. Exp. Psychol. Hum. Learn.*.

Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science, 168*(3939), 1604–1605. https://doi.org/10.1126/science.168.3939.1604.

Deutsch, D. (1975). The organization of short-term memory for a single acoustic attribute. In D. Deutsch, D., & J. A. Deutsch (Eds.). *Short-term memory* (pp. 107–151). New York: 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. 31–32). New York: Wiley.

Deutsch, D. (2013). *Grouping mechanisms in music the psychology of music* (3rd ed.). Elsevier183–248.

Deutsch, J. A., & Deutsch, D. (1963). Attention - some theoretical considerations. *Psychological Review, 70*(1), 80–90. https://doi.org/10.1037/h0039515.

Donaldson, W., & Murdock, B. B. (1968). Criterion change in continuous recognition memory. *Journal of Experimental Psychology, 76*(3P1), 325–330. https://doi.org/10.1037/H0025510.

Escudero, P., & Chládková, K. (2010). Spanish listeners' perception of American and Southern British English vowels. *Journal of the Acoustical Society of America, 128*(5), EL254–EL260.

Eysenck, M. W., & Keane, M. T. (2015). *Cognitive psychology: A student's handbook* (7th ed. ed.). London and New York: Taylor & Francis.

Fernandes, M. A., & Moscovitch, M. (2000). Divided attention and memory: Evidence of substantial interference effects at retrieval and encoding. *Journal of Experimental Psychology: General, 129*(2), 155–176. https://doi.org/10.1037/0096-3445.129.2.155.

Frauss, P. (2005). Jane austen. Retrieved from http://pia-frauss.de/fonts/ja.htm.

Friedman, D. (1990a). Cognitive event-related potential components during continuous recognition memory for pictures. *Psychophysiology, 27*(2), 136–148.

Friedman, D. (1990b). ERPS during continuous recognition memory for words. *Biological Psychology, 30*(1), 61–87. https://doi.org/10.1016/0301-0511(90)90091-A.

Gilbert, R. C., Chandrasekaran, B., & Smiljanic, R. (2014). Recognition memory in noise for speech of varying intelligibility. *Journal of the Acoustical Society of America, 135*(1), 389–399. https://doi.org/10.1121/1.4838975.

Goldstone, R. L. (1995). Effects of categorization on color perception. *Psychological Science, 6*(5), 298–304.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585–612.

Herff, S. A., & Czernochowski, D. (2019). The role of divided attention and expertise in melody recognition. *Musicae Scientiae, 23*(1), 69–86. https://doi.org/10.1177/1029864917731126.

Herff, S. A., Dean, R. T., & Olsen, K. N. (2017). Inter-rater agreement in memory for melody as a measure of listeners' similarity in music perception. *Psychomusicology: Music, Mind & Brain, 27*(4), 297–311. https://doi.org/10.1037/pmu0000197.

Herff, S. A., Dean, R. T., & Schaal, N. K. (2018). Context effects of background babbling on memory for melodies. *Musicae Scientiae.* https://doi.org/10.1177/1029864918779415 (advance online publication).

Herff, S. A., Olsen, K. N., & Dean, R. T. (2018). Resilient memories for melodies: The number of intervening melodies does not influence novel melody recognition. *Quarterly Journal of Experimental Psychology, 71*(5), https://doi.org/10.1080/17470218.2017.1318932.

Herff, S. A., Olsen, K. N., Dean, R. T., & Prince, J. (2018). Memory for melodies in unfamiliar tuning systems: Investigating effects of recency and number of intervening items. *Quarterly Journal of Experimental Psychology, 71*, 1367–1381. https://doi.org/10.1080/17470218.2017.1333519.

Herff, S. A., Olsen, K. N., Prince, J., & Dean, R. T. (2018). Interference in memory for pitch-only and rhythm-only sequences. *Musicae Scientiae, 22*, 344–361.

Hockley, W. E. (1992). Item versus associative information - further comparisons of forgetting rates. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(6), 1321–1330. https://doi.org/10.1037//0278-7393.18.6.1321.

Iidaka, T., Anderson, N. D., Kapur, S., Cabeza, R., & Craik, F. I. M. (2000). The effect of divided attention on encoding and retrieval in episodic memory revealed by positron emission tomography. *Journal of Cognitive Neuroscience, 12*(2), 267–280. https://doi.org/10.1162/089892900562093.

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*(1), 54–69. https://doi.org/10.1037/A0028347.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572.

Knoblich, G., Seigerschmidt, E., Flach, R., & Prinz, W. (2002). Authorship effects in the prediction of handwriting strokes: Evidence for action simulation during action perception. *The Quarterly Journal of Experimental Psychology: Section A, 55*(3), 1027–1046.

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*(3), 558–578. https://doi.org/10.1037/a0019165.

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

Kruschke, J. K. (2010). *Null hypothesis significance testing doing bayesian data analysis: A tutorial introduction with R.* Academic Press265–287.

Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science, 6*(3), 299–312. https://doi.org/10.1177/1745691611406925.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General, 142*(2), 573–603. https://doi.org/10.1037/a0029146.

Malmberg, K. J., & Annis, J. (2012). On the relationship between memory and perception: Sequential dependencies in recognition memory testing. *Journal of Experimental Psychology: General, 141*(2), 233–259. https://doi.org/10.1037/a0025277.

Margulis, E. H. (2005). A model of melodic expectation. *Music Perception: An Interdisciplinary Journal, 22*(4), 663–714.

Mattys, S. L., Bradlow, A. R., Davis, M. H., & Scott, S. (2013). *Speech recognition in adverse conditions: Explorations in behaviour and neuroscience.* Psychology Press.

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language & Cognitive Processes, 27*(7–8), 953–978. https://doi.org/10.1080/01690965.2012.705006.

McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review, 39*(4), 352.

Munhall, K. G., Jones, J. A., Callan, D. E., Kuratate, T., & Vatikiotis-Bateson, E. (2004). Visual prosody and speech intelligibility: Head movement improves auditory speech perception. *Psychological Science, 15*(2), 133–137.

Nathoo, F. S., & Masson, M. E. J. (2016). Bayesian alternatives to null-hypothesis significance testing for repeated-measures designs. *Journal of Mathematical Psychology,* 144–157. https://doi.org/10.1016/j.jmp.2015.03.003.

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymmetry. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(5), 1091.

Naveh-Benjamin, M., Craik, F. I., Perretta, J. G., & Tonev, S. T. (2000). The effects of divided attention on encoding and retrieval processes: The resiliency of retrieval processes. *Quarterly Journal of Experimental Psychology, 53*(3), 609–625. https://doi.org/10.1080/713755914.

Naveh-Benjamin, M., Kilb, A., & Fisher, T. (2006). Concurrent task effects on memory encoding and retrieval: Further support for an asymmetry. *Memory & Cognition, 34*(1), 90–101.

Nickerson, R. S. (1965). Short-term-memory for complex meaningful visual configurations - a demonstration of capacity. *Canadian Journal of Psychology, 19*(2), 155–160. https://doi.org/10.1037/h0082899.

Olson, G. M. (1969). Learning and retention in a continuous recognition task. *Journal of Experimental Psychology, 81*(2), 381–384. https://doi.org/10.1037/h0027756.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review, 76*(3), 241–263. https://doi.org/10.1037/h0027272.

Park, D. C., Smith, A. D., Dudley, W. N., & Lafronza, V. N. (1989). Effects of age and a divided attention task presented during encoding and retrieval on memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(6), 1185–1191. https://doi.org/10.1037/0278-7393.15.6.1185.

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. *Journal of Gerontology, 35*(1), 77–86. https://doi.org/10.1093/geronj/35.1.77.

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

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

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C., Jr. (2012a). *Word perception I: Some basic issues and methods psychology of reading.* Psychology Press49–88.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C., Jr. (2012b). *Word perception II: Word identification in text psychology of reading.* Psychology Press135–162.

Sadeh, T., Ozubko, J. D., Winocur, G., & Moscovitch, M. (2014). How we forget may depend on how we remember. *Trends in Cognitive Sciences, 18*(1), 26–36. https://doi.org/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*(1), 27–52. https://doi.org/10.1016/S0010-0277(98)00075-4.

Schellenberg, E. G. (1996). Expectancy in melody: Tests of the implication-realization model. *Cognition, 58*(1), 75–125. https://doi.org/10.1016/0010-0277(95)00665-6.

Schmidt, M. (1996). *Rey auditory verbal learning test: A handbook.* Los Angeles, CA: Western Psychological Services.

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*. New York: Springer.

Schon, D., & Francois, C. (2011). Musical expertise and statistical learning of musical and linguistic structures. *Frontiers in Psychology, 2*(167), 161–169. https://doi.org/10.3389/fpsyg.2011.00167.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*(2), 461–464.

Shepard, R. N., & Teghtsoonian, M. (1961). Retention of information under conditions approaching a steady-state. *Journal of Experimental Psychology, 62*(3), 302–309. https://doi.org/10.1037/H0048606.

Smith, S. M., & Vela, E. (1992). Environmental context-dependent eyewitness recognition. *Applied Cognitive Psychology, 6*(2), 125–139. https://doi.org/10.1002/Acp.2350060204.

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8*(2), 203–220. https://doi.org/10.3758/Bf03196157.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*(1), 34–50. https://doi.org/10.1037//0096-3445.117.1.34.

Stevens, C. J. (2015). Is memory for music special? *Memory Studies, 8*(3), 263–266. https://doi.org/10.1177/1750698015584873.

Tillmann, B., & Dowling, W. J. (2007). Memory decreases for prose, but not for poetry. *Memory & Cognition, 35*(4), 628–639. https://doi.org/10.3758/BF03193301.

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*(5), 1131–1142. https://doi.org/10.1037/0278-7393.30.5.1131.

Tse, P. U., & Cavanagh, P. (2000). Chinese and Americans see opposite apparent motions in a Chinese character. *Cognition, 74*(3), B27–B32. https://doi.org/10.1016/S0010-0277(99)00065-7.

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review, 64*(1), 49.

Wessendorf, N. (2018). *The effects of background babbling and transcranial direct current stimualtion on memory for melodies. (Bachelor of Science).* Unpublished thesis: Henrich-Heine-Universit.

Wilks, S. S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *The Annals of Mathematical Statistics, 9*(1), 60–62. https://doi.org/10.2307/2957648.