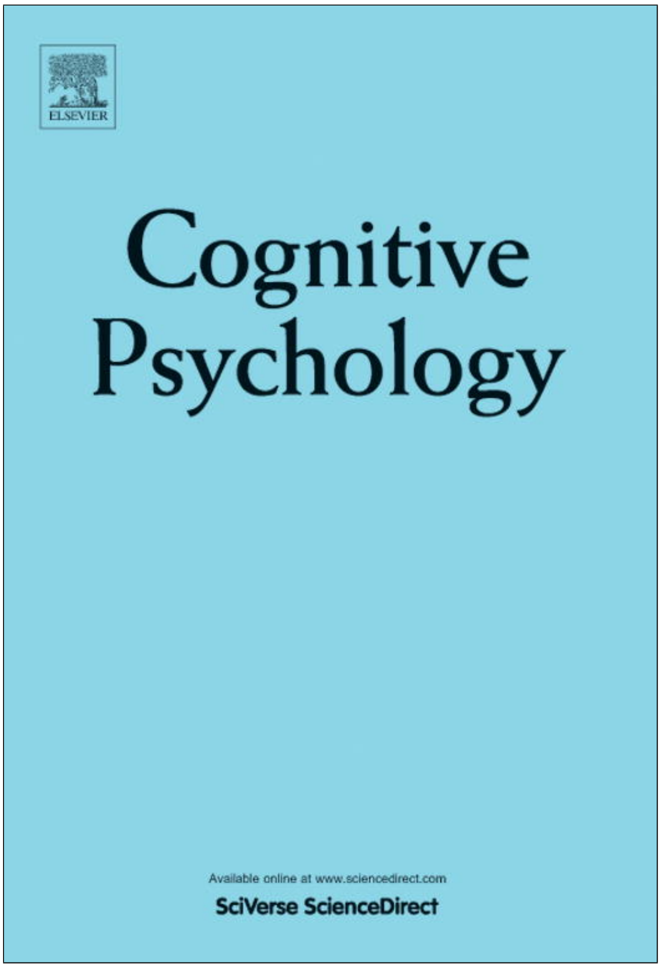
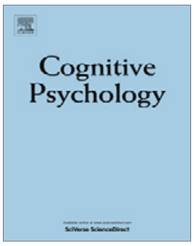

ELSEVIER
Cognitive
Psychology
Available online at www.sciencedirect.com
SciVerse ScienceDirect

Cognitive Psychology 65 (2012) 321–351

Contents lists available at SciVerse ScienceDirect

Cognitive Psychology

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



Similarity-based restoration of metrical information: Different listening experiences result in different perceptual inferences

Sarah C. Creel *

University of California, San Diego, 9500 Gilman Drive, Cognitive Science Mail Code 0515, La Jolla, CA 92093-0515, United States

A R T I C L E I N F O

Article history:
Accepted 23 April 2012

Keywords:
Perceptual restoration
Auditory learning
Meter
Time perception
Music perception
Pattern completion

A B S T R A C T

How do perceivers apply knowledge to instances they have never experienced before? On one hand, listeners might use idealized representations that do not contain specific details. On the other, they might recognize and process information based on more detailed memory representations. The current study examined the latter possibility with respect to musical meter perception, previously thought to be computed based on highly-idealized (isochronous) internal representations. In six experiments, listeners heard sets of metrically-ambiguous melodies. Each melody was played in a simultaneous musical context with unambiguous metrical cues (3/4 or 6/8). Cross-melody similarity was manipulated by pairing certain cues—timbre (musical instrument) and motif content (2–6-note patterns)—with each meter, or distributing cues across meters. After multiple exposures, listeners heard each melody without context, and judged metrical continuations (all Experiments) or familiarity (Experiments 5–6). Responses were assessed for "metrical restoration"—the tendency to make metrical judgments that fit the melody's previously-heard metrical context. Cross-melody similarity affected the presence and degree of metrical restoration, and timbre affected familiarity. Results suggest that metrical processing may be calculated based on fairly detailed representations rather than idealized isochronous pulses, and is dissociated somewhat from familiarity judgments. Implications for theories of meter perception are discussed.

---

* Fax: +1 858 534 1128.
E-mail address: creel@cogsci.ucsd.edu

http://dx.doi.org/10.1016/j.cogpsych.2012.04.004

1. Introduction

A major problem in perception and recognition is understanding how learners apply their knowledge to new instances. For example, how do we recognize a familiar song from the radio when we hear it whistled by a stranger walking down the street? Two major classes of solutions have been proposed. One is that learners determine the meaningful variation in their input and discard extraneous detail, leaving idealized representations that best capture the category. A second solution to recognition of new instances is that detail is preserved in memory, and recognition is computed based on partial activation of detailed representations. On this account, idealized representations do not exist as separate entities but emerge from consistent, recurrent patterns among experienced information. In ambiguous circumstances, consistent patterns can be filled into the perceptual experience. A variety of models exhibit this sort of pattern-completion or fill-in behavior, including neural networks (e.g. Elman & McClelland, 1988; McClelland & Elman, 1986), exemplar-style models (Goldinger, 1998; Hintzman, 1986), and Hebbian cell assemblies (Hebb, 1949). The crucial point is that, rather than an idealized representation shaping recognition and processing, multiple memories shape recognition and processing of the incoming percept in proportion to their similarity to the input.

The aim of the current study is to explore whether similarity-based processing of musical information influences percepts of musical meter, currently thought to be calculated based on highly-idealized internal representations (e.g. Large & Jones, 1999). This expands on earlier work by Creel (2011), who found that listeners hearing metrically-ambiguous melodies made metrical judgments that were consistent with the metrical contexts they had heard the melody in previously. The studies here provide a plausibility argument that meter is an emergent property of an aggregation of perceptual representations. This work also expands on demonstrations of highly-specific musical memory (e.g. Palmer, Jungers, & Jusczyk, 2001; Schellenberg & Trehub, 2003). More broadly, this study suggests a *purpose* for these specific memory representations: facilitating processing as accurately as possible.

1.1. Idealized vs. specific representations in language

Research in the domain of speech processing provides evidence both for idealized representations and for specific memories. For example, Sumner and Samuel (2005) showed that, while listeners recognized words containing non-standard (casual) pronunciations of the phoneme /t/, only words containing standard versions of /t/ showed long-term phonological priming (see also McLennan, Luce, & Charles-Luce, 2003). Relatedly, McLennan and Luce (2005) found that listeners were only sensitive to talker properties in a long-term priming experiment when their decisions were measured late in processing, implying that talker information is not central to word representations. These studies suggest that listeners possess idealized representations of speech sounds, and that non-ideal percepts feed into these representations.

Other studies, though, suggest that perceptual details are important. For instance, a memory of a particular talker saying a particular word facilitates recognition of the word (e.g. Church & Schacter, 1994; Creel, Aslin, & Tanenhaus, 2008; Creel & Tumlin, 2011; Goldinger, 1996, 1998; Palmeri, Goldinger, & Pisoni, 1993; Schacter & Church, 1992). In some studies of adaptation to atypical speech sounds, the adaptation does not transfer from one talker to another (Eisner & McQueen, 2005; Kraljic & Samuel, 2005), implying that listeners have learned talker-specific representations of these speech sounds. Lively, Logan, and Pisoni (1993; Logan, Lively, & Pisoni, 1991) find better learning of second-language speech sounds when sounds are presented from multiple talkers, suggesting that too-narrow (single-talker) representations are not sufficient for generalization (see also Clopper and Pisoni (2004) for related effects on regional-accent recognition). These studies together imply that the acoustic specifics of speech may be stored rather than discarded.[1]

---

[1] In addition to research in language, there is ample evidence that listeners store rich visual perceptual information (e.g. Brady, Konkle, Alvarez, & Oliva, 2008; Standing, 1973). Color plays a similar role in visual object recognition to talker information in word recognition: while a more classical approach suggests that achromatic edges are key to recognition (e.g. Biederman & Ju, 1988), recent work (Goffaux et al., 2005; Oliva & Schyns, 2000; Tanaka & Presnell, 1999) suggests that color facilitates recognition when it is highly diagnostic of the category. Further, Mitterer and De Ruiter (2008) have shown that object identity influences color identification. These studies are consistent with storage of detailed visual memories, with idealized object representations emerging out of the structure of the input.

1.2. Idealized vs. specific representations in music

A similar range of questions has been asked about music processing: are representations idealized or highly specific? On the one hand, listeners can identify the important pitch structures (Castellano, Bharucha, & Krumhansl, 1984; Krumhansl, Louhivuori, Toiviainen, Järvinen, & Eerola, 1999; Krumhansl et al., 2000) or metrical patterns (e.g. Ellis & Jones, 2009; Hannon, Snyder, Eerola, & Krumhansl, 2004) in unfamiliar pieces of music. That is, listeners can hear a song they have not heard before and recognize that it is in a major key. They can also recognize a familiar piece of music performed in an unfamiliar timbre, such as Happy Birthday being meowed by cats. Timbre encompasses static and dynamic acoustic attributes aside from fundamental frequency (Iverson & Krumhansl, 1993), and is not typically considered as central to musical representations as pitch or rhythm (Patel, 2008). These data are consistent with idealized representations of musical keys, meters, or melodies.

On the other hand, listeners show detailed knowledge of familiar music in terms of pitch (Levitin, 1994; Schellenberg & Trehub, 2003; Smith & Schmuckler, 2008), tempo or timing (Levitin, 1996; Palmer et al., 2001), and timbre (Gjerdingen & Perrott, 2008; Halpern & Müllensiefen, 2007; Krumhansl, 2010; Peretz, Gaudreau, & Bonnel, 1998; Poulin-Charronnat et al., 2011; Schellenberg, Iverson, & McKinnon, 1999). This sensitivity to detail raises two questions. First, does detail shape processing, or is it merely epiphenomenal? And further, if listeners have such specific knowledge, what allows them to generalize across it? The answer considered here is that similarity relationships among learned musical pieces are important in shaping processing.

1.3. Representations underlying meter perception

Musical meter—the underlying “beat” of a piece of music—is a particularly interesting case for examining similarity-based activation because meter is currently thought to be computed by comparing an incoming musical signal to relatively abstract or simple representations—internal periodic expectations. Large and Jones (1999), for example, describe meter as a periodic attentional pulse, allowing internal, periodic attending rhythms to entrain to external rhythms (see also Jones, 1976; Jones, Moynihan, Mackenzie, & Puente, 2002). Neuroimaging evidence supports activation of attending rhythms: electroencephalographic (Snyder & Large, 2005) and magnetoencephalographic (Iversen, Repp, & Patel, 2009) power oscillations are generated in certain frequency bands for musical beats that are not physically present, suggesting that listeners are implicitly predicting a metrical continuation—the next “beat.” Many authors have assessed the role of various factors in the musical surface which activate idealized metrical representations, including temporal accents (note duration) and melodic accents (changes in pitch direction; Ellis & Jones, 2009; Essens & Povel, 1985; Hannon et al., 2004; Povel & Essens, 1985).

Is there a strong role for detailed memory in meter perception? On the one hand, perhaps not—perhaps humans are simply biased to attend to periodicity, as it is so basic to biologically-generated events (e.g. Cummins & Port, 1998; Cutler & Butterfield, 1992; Dilley & McAuley, 2008; Saygin, Driver, & de Sa, 2008). Further, current approaches (discussed above) make a strong case that surface cues activate very basic internal representations to generate meter perception. Nonetheless, different experiences shape meter perception. Adults exposed to different musical cultures interpret metrical patterns differently (Hannon & Trehub, 2005a), and only Western 12-month-olds—not Western adults—can process culturally-unfamiliar metrical patterns after just 2 weeks of exposure (Hannon & Trehub, 2005b). Language or culture differences influence interpretation of ambiguous strong–weak/weak–strong beat patterns (Iversen, Patel, & Ohgushi, 2008; Yoshida et al., 2010). Finally, movement experience may shift interpretation of an ambiguous meter in infants and adults (Phillips-Silver & Trainor, 2005, 2007, 2008).[2] These studies together suggest that meter may rely upon perceptual representations, possibly accumulated over a lengthy time span, that are more detailed than isochronous pulses.

---

[2] However, McAuley (2011) and McAuley, Henry, Rajarajan, and Nave (2011) make a strong case that the effect is not about rhythm perception but about participants' interpretations of the experimenter's expectations.

There are several ways in which memory representations more specific than periodic oscillations might provide a broader explanatory framework for metrical processing. First, a memory-based explanation may provide a better account of "complex" (non-isochronous) meters (Hannon & Trehub, 2005a, 2005b) found in a number of cultures. Such meters are easily recognized and processed by enculturated listeners (Hannon & Trehub, 2005a). Periodic oscillators (Large & Jones, 1999) would have difficulty capturing this behavior (see London (1995) for supporting arguments). If oscillators were replaced with not-necessarily-isochronous memory patterns, then coactivated similar memories (in enculturated listeners) would simply "play along" with the percept, providing support for the metrical pattern.

Second, specific memories might aid in recognition (e.g. Gjerdingen & Perrott, 2008; Krumhansl, 2010) and processing of differing musical styles. For instance, Western classical music and American jazz music contain different timbres, melodic figures, harmonic progressions, and metrical or rhythmic patterns. Classical music rarely employs the alto saxophone or swing rhythms (George Gershwin notwithstanding), but jazz music regularly employs both. If hearing an alto saxophone or a jazz-typical chord selectively activated "swing" rhythm representations more than classical-music rhythm representations, processing of the swing rhythm would be facilitated.[3] Current models of metrical processing do not incorporate this information, and thus cannot benefit from it.

Finally, influences of similarity-based memory activation on meter constitute an argument about the nature of meter itself: perhaps it is not the simple recognition of isochrony, but the coactivation of similar memory representations. The assertion of the current study is that previous accounts of meter are incomplete without taking into account activation of more detailed memory information. Specifically, I argue that meter perception is an *interaction* between surface cues in the immediate percept and multiple activated memories of similar musical information.

1.4. The current study

The current study asked whether similarity *across* melodies influences metrical processing, and what counts as similar. Effects of cross-melody similarity on metrical restoration make a plausibility case that musical processing arises out of an aggregate of musical memories, rather than from idealized representations. To examine this possibility, the study manipulates participants' percepts of the 3/4–6/8 metrical ambiguity. This ambiguity might be thought of as a "Necker cube" of event timing (Fig. 1): it can be organized as three groups of two sub-beats, or as two groups of three sub-beats (see Handel & Lawson, 1983; Hannon et al., 2004). Many artists have played on this ambiguity, including Antonín Dvořák (*Slavonic Dances*), Johannes Brahms (*Symphony No. 3*), and Tears for Fears ("Everybody Wants to Rule the World"). Metrical importance in Fig. 1 is marked as *x*'s, following Lerdahl and Jackendoff's (1983) metrical grid. The onset of each group of two or three (the second level of ×'s in Fig. 1) is the tactus—the "beat," the frequency at which listeners would tap their feet or clap their hands in time to the music. Onset events are perceived as stronger than the other sub-beat events.

Recently, Creel (2011) demonstrated that melodies are stored along with metrical information in memory, referred to in the current paper as "metrical restoration:" listeners form specific memories of melodies that shape processing of meter when the melodies are heard later. This is perhaps most analogous to continuity phenomena such as the phonemic restoration effect (Samuel, 1981; Warren, 1970), where, *on a word-specific basis*, listeners perceive the presence of a phoneme (or tone; DeWitt & Samuel, 1990) that has been completely replaced by noise (as in *dino*∗*aur*). Illusory metrical patterns have been noted in musical (Brochard, Abecasis, Potter, Ragot, & Drake, 2003; Iversen et al., 2008) and linguistic (Pitt & Samuel, 1990) meter previously, suggesting that perception of loudness alternation or emphasis alternation is subject to noise within the perceptual system or in memory. However, these previous metrical demonstrations were not shown on an instance-specific basis.[4]

In Creel's (2011) Experiment 2, listeners heard melodies that were ambiguous between 3/4 meters and 6/8 meters. Listeners heard an exposure phase and a following probe (test) phase. In exposure, a

---

[3] This is perhaps even more true of traditional musical systems *across* cultures.

[4] Further analogies include top-down perception phenomena such as sine-wave speech analogs, which listeners perceive and understand as speech when informed as to its nature (Remez, Rubin, Pisoni, & Carrell, 1981), and recognition of melodies with notes' octaves changed when informed as to the melody's identity (Deutsch, 1972).

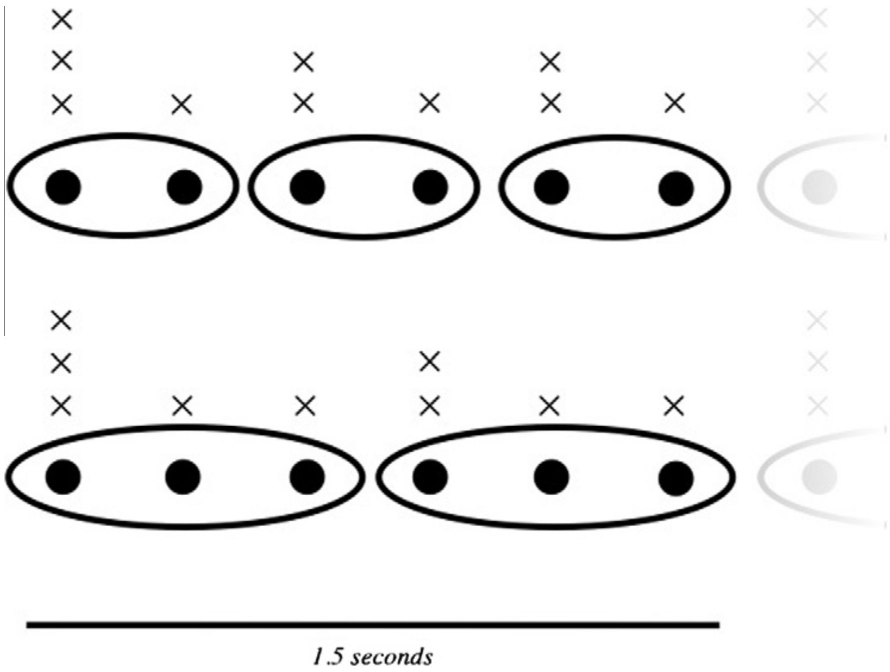


Fig. 1. Example of alternative groupings of isochronous pulses (black dots): into three groups of two (3/4 meter, top line); or two groups of three (6/8 m, bottom line). ×'s mark a metrical grid (Lerdahl & Jackendoff, 1983), with more *x*'s having greater emphasis.

clear sense of the beat for each melody was provided by *context* instruments that played simultaneously with the ambiguous melody. After multiple in-context exposures, listeners proceeded to the probe phase where they heard the melodies out of context, and judged how well 3/4 vs. 6/8 metrical probes fit those melodies. Listeners preferred the meter they had heard with that melody during exposure, implying that hearing the melody alone reactivated metrical properties of the original context. Thus an exact replication of a melody, preserving timbre, pitch, and timing, resulted in metrical restoration. However, it was not clear how, or even if, these effects might generalize to new musical material, or how they might reflect the nature of musical knowledge as a whole.

The current study used Creel's (2011) paradigm, but manipulated similarity amongst the melodies heard by a given listener, to assess the effects of cross-melody similarity on metrical processing. The question throughout is whether metrical restoration of one melody is *strengthened* by similar melodies having the same meter, or *weakened* by similar melodies having different meters. Experiments were divided into two major sections. In Part I, *timbre similarity* across melodies was manipulated to determine the influence of timbre on similarity-based activation in metrical restoration. Would a change in timbre at test result in loss of the metrical restoration effect? Further, was metrical restoration stronger when each meter had a consistent timbre? In Part II, an additional element of similarity was added: motifs. Listeners were exposed to non-identical melodies composed of two different sets of motifs—pitch-and-rhythm "chunks." Do non-identical melodies composed of identical motifs coactivate each other in memory, such that metrical restoration of one melody is stronger when other melodies with the same motifs have the same meter, and weaker when melodies with those motifs have different meters? And is metrical restoration predicted simply by recognizing the melody, or by activation of similar melodies apart from explicit recognition?

2. Part I

2.1. Experiment 1: pairing timbre with meter (A)

This initial experiment explored metrical restoration when listeners were exposed to melodies that were fairly similar to one another. This replicated Creel's (2011) earlier study by verifying whether listeners were able to form melody-specific knowledge of meter, and extended it by examining the effects of inter-melody similarity. While Creel used melodies that varied widely in timbre, presentation rate, and pitch collections, the current set of melodies were matched on presentation rate and pitch collections, with only two timbres used. For a given listener, each timbre was paired with just one meter—for instance, all saxophone melodies were heard in 3/4 contexts, and all French horn melodies, in 6/8. After exposure to melodies in metrically-clear contexts, listeners then heard melodies without their contexts, followed by metrical probes in 3/4 or in 6/8.

On an idealized-meter-representations account, meter perception for contextless melodies would be driven only by cues in the melodies themselves (and listeners' own idiosyncratic biases), not by previous listening experiences. Thus, no effects of pre-exposure on metrical processing should be evident. However, on a similarity-based learning account, timbre should influence grouping of traces in memory: hearing all saxophone melodies presented in (for example) 3/4 meters should create strong associations between the saxophone timbre and 3/4 meters. Thus, when a saxophone melody was heard alone, listeners should restore the 3/4 meters readily (and vice versa for French horn melodies and 6/8 meters). What about the same melody with a change in timbre? If timbre was strongly associated with the meter, then when a familiarized saxophone melody was heard in a different timbre, the 3/4 metrical preference should weaken, or even reverse if it is heard in the timbre associated with 6/8 meters.

2.1.1. Method

2.1.1.1. Participants. $N = 36$ college-aged participants from the UCSD psychology participant pool received course credit for participation. None in any experiment reported hearing problems. Seven additional participants were replaced due to a programming error which resulted in the loss of 40–85% of test trials. Three of the included participants were also affected by this error, but their data were retained as they were mostly intact (13% or fewer trials lost). In this and following experiments, incomplete adherence to lab procedure meant that music background data were not collected from a minority of participants (about one in five). Generally speaking, the sample was quite varied in musical sophistication, both in terms of playing music and taking music coursework. Effects of music experience, reported in Appendix A, on the experimental manipulations were small and usually not statistically significant.

2.1.1.2. Stimuli. Eighteen major-key melodies (Fig. 2) were composed in Finale 2009.r2 for Macintosh (2008, MakeMusic Inc.). Melodies were composed using a limited vocabulary of what the author judged to be moderately-metrically-ambiguous measure-long rhythmic patterns (Appendix B). Each melody was roughly eight measures in length plus a final note, with an eighth-note (sub-beat) duration of 250 ms. They were distributed across a variety of different keys (Ab major, Bb major, C major, D major, E major, F# major), so that three melodies each were in the same key. The two timbres (selected from Finale's SmartMusic SoftSynth MIDI collection) were soprano saxophone with harp accompanying, and French horn with accordion and acoustic grand piano accompaniment. The accompanying contexts (hereafter, just *contexts*) were composed so as to be metrically unambiguous. Metrical clarity was achieved by placing more note onsets on strong beats in 3/4 or 6/8 than on weak beats (as in Fig. 2a), by presenting pattern repetitions (see Hannon et al., 2004) at the length of a beat (as in Fig. 2b), or both. In-context and isolated melodies were exported as .aiff files, and were edited in Praat (Boersma & Weenink, 2007) to equalize loudness to 70 dB SPL.

2.1.1.3. Procedure. The experiment had an *exposure phase* followed by a *probe phase*. In exposure, each listener heard 12 of the 18 melodies. Six melodies were heard in one timbre + meter combination, and six others were heard in another timbre + meter combination. Each melody was heard six times each, for a total of 72 exposure trials. Each melody was equally likely to occur in each of the four timbre + meter combinations, though a given participant would only hear that particular melody in one of them (see Table 1 for counterbalancing details). So as not to alert listeners to the nature of the upcoming task, they were simply asked to rate liking and emotional valence (happy/sad) for each occurrence of each melody, as in Creel (2011).

In the probe phase, all participants heard the same materials, meaning that different responding could only be due to prior exposure (and, as in any experiment, participant-specific idiosyncrasies).

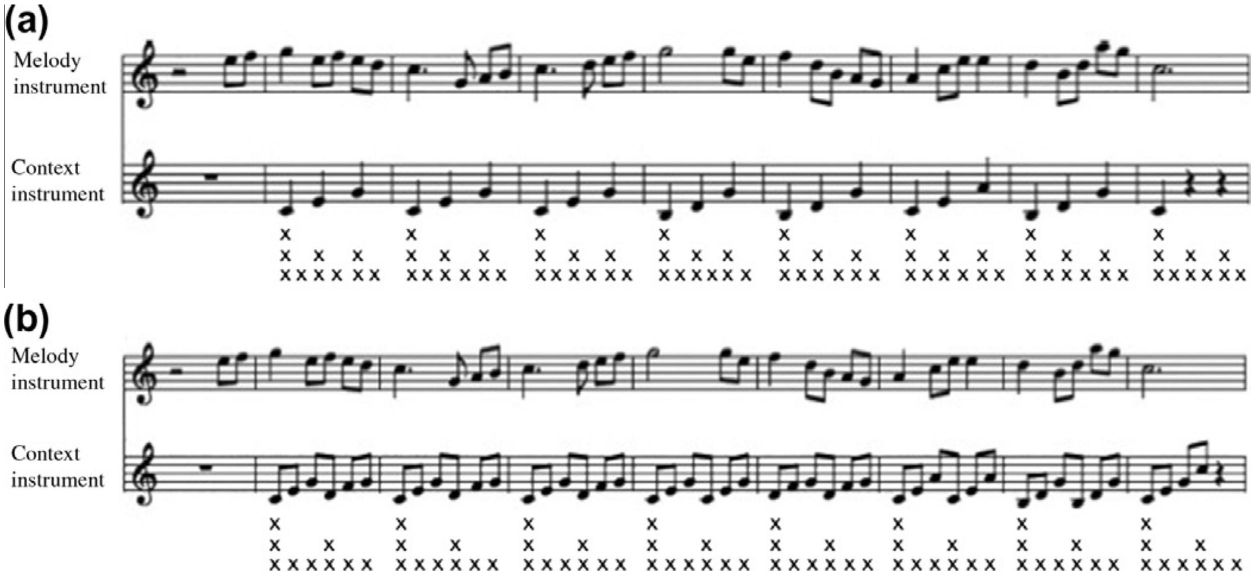



Fig. 2. Example of a melody used in Part I (Experiments 1–3). (a) in a 3/4 context, and (b) in a 6/8 context. Metrical grids for each context are marked in *x*'s below the lower staff.

Table 1
Experiments 1–2, counterbalancing of melody exposure. There were 12 lists total. Lists 7–12 were identical to Lists 1–6 except that timbres were reversed.

Melody	List 1	List 2	List 3	List 4	List 5	List 6
6	Hn, 3/4	.	Sx, 6/8	Sx, 6/8	.	Hn, 3/4
8	Hn, 3/4	.	Sx, 6/8	Sx, 6/8	.	Hn, 3/4
18	Hn, 3/4	.	Sx, 6/8	Sx, 6/8	.	Hn, 3/4
14	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8	.
12	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8	.
10	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8	.
16	.	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8
2	.	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8
4	.	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8
11	Hn, 3/4	.	Sx, 6/8	Sx, 6/8	.	Hn, 3/4
5	Hn, 3/4	.	Sx, 6/8	Sx, 6/8	.	Hn, 3/4
13	Hn, 3/4	.	Sx, 6/8	Sx, 6/8	.	Hn, 3/4
9	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8	.
3	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8	.
7	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8	.
17	.	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8
1	.	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8
15	.	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8

Note. Hn = French horn; Sx = soprano saxophone; . = did not hear the melody during exposure.

Each melody was stripped of its context and was presented in one of three timbres (saxophone, French horn, or a completely new timbre, the dulcimer). Each melody was followed by one of two probes (3/4 metrical probe or 6/8 metrical probe). Each probe consisted of four measures of beats after the end of the melody, plus the first beat of the fifth measure, and were played in a woodblock timbre. Thus, each of the 18 melodies was heard 6 times (3 timbres $*$ 2 probes), for a total of 108 test trials. This meant that each listener heard:

- the 12 familiar melodies in the original timbre;
- the 12 familiar melodies, switched to the other exposed timbre;
- the 12 familiar melodies, switched to the completely-novel timbre;

- the 600 melodies in the 3/4-associated timbre;
- the 600 melodies in the 6/8-associated timbre;
- the 600 melodies in the novel timbre.

On each test trial, the listener saw a 400-pixel-wide ruler, which was labeled "Bad fit" on the left end, "Okay" in the center, and "Good fit" on the right end. They were told "After each song, you will hear some drumbeats. Your job is to judge, on the 'ruler' below, how well the drumbeats went with the previous song. Just click on the point on the ruler that seems to best describe how well it fit. Try to use the whole range of the ruler." After reading the instructions, they were allowed to click on-screen buttons to play four examples: one phrase (four measures) of Happy Birthday (which is in 3/4 time) followed by "good drumbeats" (in 3/4); one phrase Happy Birthday followed by "bad drumbeats" (in 6/8); one phrase of Greensleeves (which is in 6/8 time) followed by "good drumbeats" (6/8); and one phrase of Greensleeves followed by "bad drumbeats" (3/4). The *x*-coordinate of each response click was transformed so that responses ranged from -1 to +1.

2.1.2. Results

Metrical probe ratings (Fig. 3a) suggested that exposed melodies showed metrical restoration effects when presented in the original timbre. Specifically, when listeners heard melodies presented in their original timbres (left four bars of Fig. 3a), ratings of 3/4 metrical probes were much higher than ratings of 6/8 probes for 3/4-exposed melodies, but 3/4 ratings were *not* higher than 6/8 ratings for 6/8-exposed melodies. An ANOVA on probe ratings for exposed melodies was conducted with Learned On (horn = 3/4/sax = 6/8, sax = 3/4/horn = 6/8) as a between-participants factor, and Exposure Meter (exposed 3/4, exposed 6/8), Probe Meter (3/4 probe, 6/8 probe), and Timbre Match (original, other [switched timbre], novel) as within-participants and within-items factors. Effect size is reported as generalized eta squared (η^2_G; Bakeman, 2005; Olejnik & Algina, 2003), which is preferable to partial eta squared because it is comparable across between-participants and within-participants designs.

If metrical restoration occurred, as in Creel (2011, Experiment 2), there should be a significant interaction between Exposure Meter and Probe Meter. This would indicate that listeners rated the two probe meters differently depending upon the exposure meter. In the current experiment, the Exposure Meter × Probe Meter interaction was significant ($F1(1,34) = 7.99$, $p = .008$; $F2(1,17) = 3.22$, $p = .09$; $\eta^2_G = .35$): listeners preferred metrical probes that matched the contexts they had previously heard with particular melodies. There was also an effect of Timbre Match × Probe Meter ($F1(2,68) = 3.62$, $p = .03$; $F2(2,34) = 4.15$, $p = .02$; $\eta^2_G = .10$), resulting from similar 6/8 probe ratings across timbre match conditions, but differing 3/4 ratings, with the largest in the novel timbre case. No other effects or interactions approached significance, including a Timbre Match × Exposure Meter × Probe Meter interaction, which would have indicated differing patterns of meter preference across changes in timbre ($F1(2,68) = 1.96$, $p = .15$; $F2(2,34) = 2.01$, $p = .15$; $\eta^2_G = .07$).

Planned analyses of each level of Timbre Match examined the Exposure Meter × Probe Meter interaction. Only the original-timbre trials showed the signature interaction pattern ($F1(1,35) = 12.11$, $p = .001$; $F2(1,17) = 6.26$, $p = .02$; $\eta^2_G = .59$), while the changed-timbre trials ($F1(1,35) = 1.36$, $p = .26$; $F2(1,17) = 1.86$, $p = .19$; $\eta^2_G = .17$) and the novel-timbre trials ($F1(1,35) = 1.78$, $p = .19$; $F2 < 1$; $\eta^2_G = .16$) did not. Note that a simpler way to depict the results is to plot the difference between matched probes and mismatched probes: (exposed 3/4, 3/4 probe + exposed 6/8, 6/8 probe)/2 − (exposed 3/4, 6/8 probe + exposed 6/8, 3/4 probe)/2, as in Fig. 3c. This is directly proportional to the interaction score: (exposed 3/4, 3/4 probe − exposed 3/4, 6/8 probe) − (exposed 6/8, 3/4 probe − exposed 6/8, 6/8 probe). When this difference score is zero, there is no metrical restoration; when it is greater than zero, there is metrical restoration. For ease of interpretation, results for following experiments are plotted in this way. Means for individual probe meter × exposure meter cells are provided in Appendix C (Tables C1–C6).

To see if timbre-related meter preferences carried over to unheard melodies (Fig. 3b), ANOVAs were conducted with Learned On, Exposure Meter (timbre heard with 3/4 during exposure, timbre heard with 6/8 during exposure) and Probe Meter as within-participants and within-items factors. There was an effect of Probe Meter ($F1(1,34) = 4.61$, $p = .04$; $F2(1,17) = 13.55$, $p = .002$; $\eta^2_G = .75$), suggesting an overall preference for 3/4 probes, but no other effects approached significance. There was

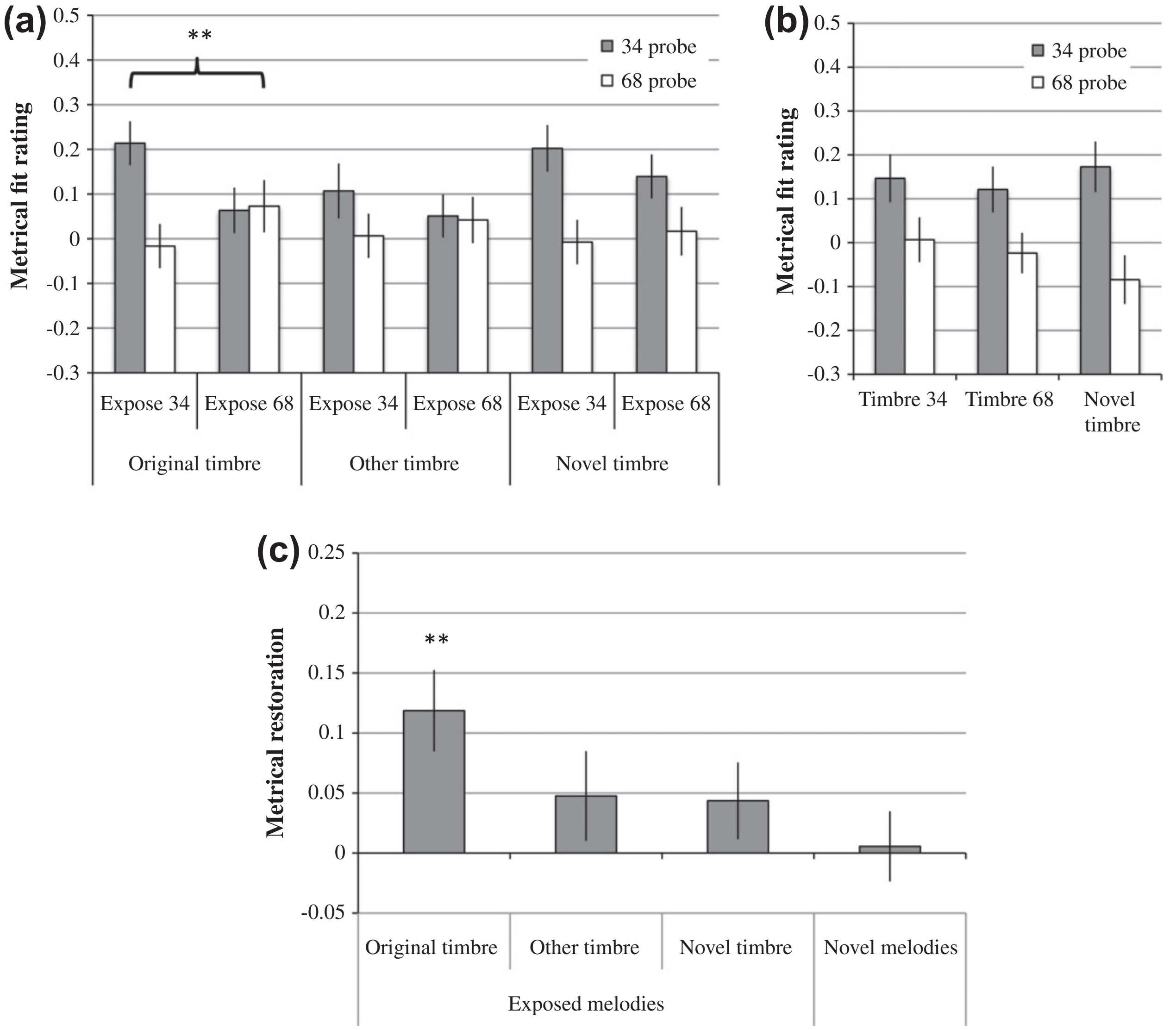


Fig. 3. Experiment 1, results for (a) familiar-melody trials and (b) novel-melody trials; (c) simplified display of the data in (a) and (b), as a difference score between exposure matching-probes vs. exposure-mismatching probes (see text for details). Throughout, error bars are standard errors.

no strong prediction for novel-timbre unheard melodies; analysis suggested a strong 3/4 preference ($t1(35) = 2.81$, $p = .008$; $t2(17) = 4.21$, $p = .0006$; 3/4: $M = .20 \pm .32$; 6/8: $M = -.06 \pm .33$).

2.1.3. Discussion

This initial experiment explored whether cross-melody similarity in timbre influenced metrical restoration. Listeners heard sets of melodies with distinct timbre and meter combinations. Listeners showed melody-specific learning effects for meter, and these effects were strongest when the melody was heard in its original timbre. However, there was not a statistical interaction of Timbre × Exposure Meter × Probe Meter, making it difficult to conclude whether timbre truly mattered for metrical restoration.

In addition to the metrical learning effect, there was a bias toward 3/4 probes. This was observed in Creel's (2011, Experiment 2) control condition: without exposure, listeners strongly preferred to interpret ambiguous melodies as being in 3/4. In particular, the cases in the current experiment where the strongest 3/4 biases were observed were those least similar to the exposure (novel melodies, esp. when played by the novel instrument). This suggests that learning effects interact with preexisting biases, such as preferences for certain tempos (e.g. London, 1995).

Interestingly, the magnitude of the metrical-restoration effect—a rating difference between exposure-matching and exposure mismatching meters—is only .12 ($SD = .20$), even in the original-timbre condition. This is much smaller than that observed in Creel (2011), with a rating difference of .56 ($SD = .43$). This may be due to the tighter similarity structure among the current set of melodies, which had two timbres (Creel used 6), a single rate of presentation (Creel used 3), and a single pitch collection (major-key; Creel used four different scales). A reduction in the metrical restoration effect as musical memories become more similar is consistent with a similarity-based memory account, as more similar memories will interfere more with each other.

Effects in the current experiment may also have been weakened by the lengthy probe portion of the experiment: at 108 trials, it lasted longer than the exposure phase (72 trials), which may have diluted the short-term learning during exposure. Given that Experiment 1 hinted at timbre-specificity of metrical restoration, the next experiment replicated it, but with a more sensitive (shorter) test phase.

2.2. Experiment 2: pairing timbre with meter (B)

2.2.1. Method

2.2.1.1. Participants. $N = 36$ new participants from the same pool as in Experiment 1 took part.

2.2.1.2. Stimuli and Procedure. These were the same as in Experiment 1, except that novel-timbre versions of melodies were removed from the probe phase.

2.2.2. Results

With a smaller number of trials, the suggestion of timbre-specific meter learning in Experiment 1 was replicated with a significant Timbre Match × Exposure Meter × Probe Meter interaction (Fig. 4; note qualitative similarity to Fig. 3c). An ANOVA was conducted on metrical probe ratings with Exposure Meter, Probe Meter, and Timbre Match as within-participants factors, and Learned On as a between-participants factor. The interaction of Exposure Meter × Probe Meter was significant ($F1(1,34) = 9.03$, $p = .005$; $F2(1,17) = 7.26$, $p = .02$; $\eta^2_G = .38$), suggesting that 3/4 probes were rated higher after 3/4-exposed melodies than after 6/8-melodies, and vice versa. There was also an interaction of Timbre Match × Exposure Meter × Probe Meter ($F1(1,34) = 15.08$, $p = .0005$; $F2(1,17) = 16.45$, $p = .0008$; $\eta^2_G = .32$). Simple ANOVAs showed that only original-timbre trials showed the Exposure

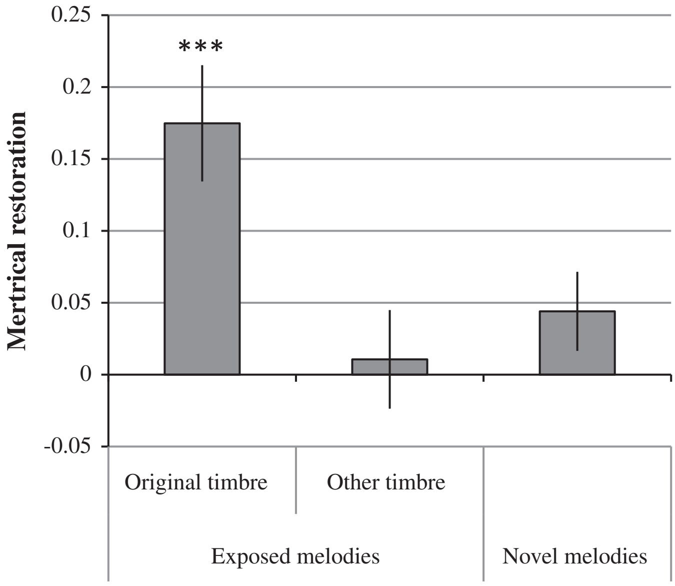


Fig. 4. Experiment 2, metrical fit ratings plotted as match–mismatch difference scores to better illustrate degree of metrical restoration.

Meter × Probe Meter interaction ($F1(1,35) = 18.2$, $p = .0001$; $F2(1,17) = 20.57$, $p = .0003$; $\eta^2_G = .67$), while other-timbre trials did not ($F < 1$; $F2 < 1$; $\eta^2_G = .008$).

For unheard melodies, ANOVAs were conducted with Learned On, Timbre (timbre with 3/4 exposure, timbre with 6/8 exposure) and Probe Meter as within-participants and within-items factors. No effects approached significance, including the Timbre × Probe Meter interaction ($F1(1,34) = 2.45$, $p = .13$; $F2(1,17) = 1.99$, $p = .18$; $\eta^2_G = .13$), suggesting that metrical preferences did not strongly transfer to new melodies based on timbre similarity.

2.2.3. Discussion

Listeners were shifted toward metrical fit judgments of melodies that matched with their experience of those melodies and their specific timbres. These data suggest that listeners group musical information in memory by timbral similarity: the consistent linkage between timbre and meter strengthened timbre–meter connections, such that metrical restoration only occurred for a melody heard in its original timbre. This bolsters the equivocal evidence for timbre specificity in Experiment 1: timbre similarity does influence metrical restoration. Like Experiment 1, this did not transfer to novel melodies in the exposed timbres.

This experiment suggests that coherence amongst multiple representations in timbre and meter leads to timbre-specific restoration of metrical information. However, there is an alternative explanation, which has nothing to do with timbre similarity across melodies: listeners may simply be storing *each melody* in timbre-specific form, without any interaction in memory between same-timbre melodies. If that is the case, then metrical preferences for individual melodies should show timbre-specificity—a drop in metrical restoration when the melody's timbre changes at test—even if there is no experiment-wide consistency in timbre–meter mapping. This was tested in Experiment 3.

2.3. Experiment 3: timbre associated with melodies but not meters

2.3.1. Method

2.3.1.1. Participants. $N = 36$ new participants from the same pool as in Experiments 1 and 2 took part.

2.3.1.2. Stimuli and procedure. These were the same as the two previous experiments, except that exposure for each participant was rearranged (Table 2) such that each person heard three melodies each with French horn +3/4, French horn +6/8, saxophone +3/4, and saxophone +6/8, for a total of 12 exposure melodies.

2.3.2. Results

Listeners showed melody-specific meter preferences, but these preferences generalized across changes in timbre (Fig. 5). An ANOVA on metrical fit ratings was conducted with Exposure Meter, Probe Meter, and Timbre Match (original, other) as within-participants factors. Probe Meter was significant ($F1(1,35) = 5.64$, $p = .02$; $F2(1,17) = 10.28$, $p = .005$; $\eta^2_G = .78$), indicating an overall preference for 3/4 probes. The Exposure Meter × Probe Meter interaction was also significant ($F1(1,35) = 13.12$, $p = .0009$; $F2(1,17) = 4.86$, $p = .04$; $\eta^2_G = .51$), indicating that listeners rated probes higher when they matched the previously-heard context's meter. However, the higher-level interaction of Timbre Match × Exposure Meter × Probe Meter ($F1 < 1$; $F2 < 1$; $\eta^2_G = .01$) was not significant, suggesting that effects of metrical exposure were similar regardless of whether the melody was played in the original timbre or the other timbre. Novel melodies showed a marginal trend toward preference of 3/4 probes ($t1(35) = 1.77$, $p = .09$; $t2(17) = 1.58$, $p = .13$).

2.3.3. Discussion

Listeners in Experiment 3 heard melodies of two different timbres, but the timbres did not map to particular meters. The listeners showed a metrical restoration effect, but, unlike the first two experiments, they did not show timbre-specific metrical knowledge. That is, metrical restoration was tied to each melody, but not to timbre. This means that timbre-specific metrical restoration in Experiments 1 and 2 cannot be explained as resulting from individual-melody knowledge. Instead, listeners in the first two experiments seem to have formed memory collections that linked together melodies of

Table 2
Experiment 3, counterbalancing of melody exposure. There were 18 lists total. Lists 10–18 were identical to Lists 1–9 except that timbres were reversed.

Melody	List 1	List 2	List 3	List 4	List 5	List 6	List 7	List 8	List 9
6	Hn, 3/4	.	Sx, 6/8	Sx, 3/4	.	Hn, 6/8	Sx, 6/8	.	Hn, 3/4
8	Hn, 6/8	.	Sx, 3/4	Sx, 6/8	.	Hn, 3/4	Sx, 6/8	.	Hn, 3/4
18	Hn, 3/4	.	Sx, 6/8	Sx, 6/8	.	Hn, 3/4	Sx, 3/4	.	Hn, 6/8
14	Sx, 6/8	Hn, 3/4	.	Hn, 6/8	Sx, 3/4	.	Hn, 3/4	Sx, 6/8	.
12	Sx, 3/4	Hn, 6/8	.	Hn, 3/4	Sx, 6/8	.	Hn, 3/4	Sx, 6/8	.
10	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8	.	Hn, 6/8	Sx, 3/4	.
16	.	Sx, 6/8	Hn, 3/4	.	Hn, 6/8	Sx, 3/4	.	Hn, 3/4	Sx, 6/8
2	.	Sx, 3/4	Hn, 6/8	.	Hn, 3/4	Sx, 6/8	.	Hn, 3/4	Sx, 6/8
4	.	Sx, 6/8	Hn, 3/4	.	Hn, 3/4	Sx, 6/8	.	Hn, 6/8	Sx, 3/4
11	Hn, 6/8	.	Sx, 3/4	Sx, 6/8	.	Hn, 3/4	Sx, 3/4	.	Hn, 6/8
5	Hn, 3/4	.	Sx, 6/8	Sx, 3/4	.	Hn, 6/8	Sx, 3/4	.	Hn, 6/8
13	Hn, 6/8	.	Sx, 3/4	Sx, 3/4	.	Hn, 6/8	Sx, 6/8	.	Hn, 3/4
9	Sx, 3/4	Hn, 6/8	.	Hn, 3/4	Sx, 6/8	.	Hn, 6/8	Sx, 3/4	.
3	Sx, 6/8	Hn, 3/4	.	Hn, 6/8	Sx, 3/4	.	Hn, 6/8	Sx, 3/4	.
7	Sx, 3/4	Hn, 6/8	.	Hn, 6/8	Sx, 3/4	.	Hn, 3/4	Sx, 6/8	.
17	.	Sx, 3/4	Hn, 6/8	.	Hn, 3/4	Sx, 6/8	.	Hn, 6/8	Sx, 3/4
1	.	Sx, 6/8	Hn, 3/4	.	Hn, 6/8	Sx, 3/4	.	Hn, 6/8	Sx, 3/4
15	.	Sx, 3/4	Hn, 6/8	.	Hn, 6/8	Sx, 3/4	.	Hn, 3/4	Sx, 6/8

Note. Hn = French horn; Sx = soprano saxophone; . = did not hear the melody during exposure.

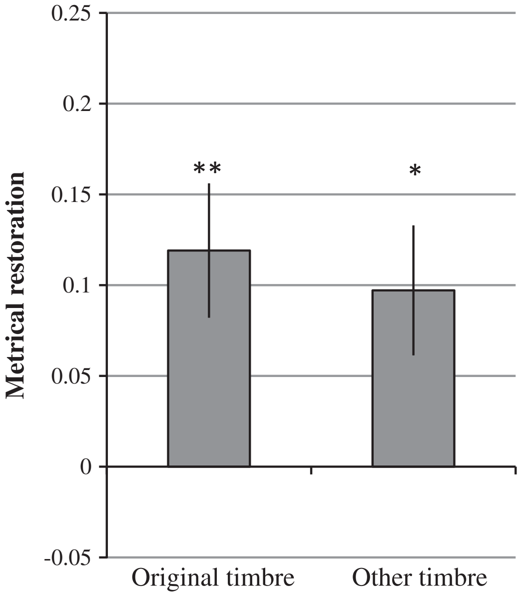


Fig. 5. Experiment 3, metrical fit ratings plotted as match–mismatch difference scores.

the same timbre. It also suggests that, when not consistently associated with meter, timbre plays little role in activating metrical memory information.

2.4. Interim summary

Listeners are sensitive to timbre–meter coherence in their musical experiences. Specifically, listeners show metrical restoration effects for previously-heard melodies when they hear the melodies again *in the original timbre*, but not when they hear them in a different timbre. When timbres are equally likely to occur with either meter, listeners still show metrical restoration, but timbre does not matter. Results do not suggest a simple association between timbre and meter—if that were the case, then listeners in Experiments 1–2 should have preferred a particular meter for *all* melodies of one timbre—both familiarized and novel ones—not just the familiar melodies. What the results suggest instead is that similar memories bolster each other. This suggests not only that listeners recognize music based on timbral detail (Halpern & Müllensiefen, 2007; Schellenberg et al., 1999), but that timbre similarity across multiple melodies can influence their perception of the music's meter. This is interesting in that timbre is often construed as a less-structurally-essential property of music than pitch or rhythm (Patel, 2008). However, as noted in the Introduction, different musical styles utilize both characteristic metrical patterns *and* characteristic instrumental timbres, and such a timbre-specific memory activation mechanism might facilitate recognition and metrical processing in a variety of styles.

These experiments leave certain questions unanswered, such as why listeners do not extend metrical restoration to novel melodies with the trained timbres. This raises the related question of what properties of the melodies contribute to metrical restoration. Must listeners recognize the melody in order to restore meter, with timbre match increasing the likelihood of recognition? If so, then the strength of metrical restoration should vary directly with familiarity. On the other hand, listeners may not be remembering melody–meter pairings wholesale, but may instead be attaching metrical information to subparts of the melodies. The latter would predict that rearranging motifs from a variety of 3/4-exposed melodies would result in a 3/4-like percept, even if the melody is not recognized as familiar. Broadly speaking, a *motif* is a brief, recurring "chunk" (two or more notes, but rarely an entire musical phrase) used throughout a musical composition (Randel, 1986). If listeners are storing metrical information in relation to some melodic subpart, motifs are a likely candidate. Part II explores the influence of motif content in metrical restoration.

3. Part II

For the second set of studies, two new sets of melodies were constructed from two sets of eight measure-long motifs (Appendix D). Motifs were randomly assigned to the two sets. Each motif was a rhythmic pattern with a particular contour (patterns of ups and downs in pitch). However, the exact size of pitch changes between successive notes in a given motif was allowed to vary. Each melody was a unique ordering of the eight motifs in its set. Note that there is some rhythmic similarity *within* the motifs in a set (e.g. Set B has three 1 1 1 1 1 1 rhythms), but there is also similarity in motifs *across* sets (both Set A and Set B have 3 1 2 rhythms and 3 1 1 1 rhythms). Sets of motifs are nearly identical in average number of notes per motif (Set A: 4.13; Set B: 4.25). The reader may observe that Part I also used a small set of rhythms. However, counting these using the definition of motif employed in Part II shows *54* distinct rhythm + contour patterns, far more than the 16 used in Part II. Thus, Part II melodies have tighter similarity structures. On a similarity-based memory activation account, this tighter similarity structure may both create stronger metrical restoration effects, and interfere with metrical restoration if motifs are scattered across meters. As before, timbre match was also manipulated.

3.1. Experiment 4: probing initial metrical experiences

Before describing manipulations of timbre and motifs in learning, naïve responding is characterized. Without hearing an exposure phase, listeners were presented with a probe phase and asked to judge metrical fit. Some listeners heard all melodies in all possible meter and timbre contexts, to judge the effectiveness of contexts at eliciting the intended metrical percept at the time of first hearing. Other listeners heard all melodies in both timbres, but without context, to judge the baseline metrical percepts of melodies. These data established baseline outcomes for the melodies used in Part II.

3.1.1. Method

3.1.1.1. Participants. There were $N = 36$ participants from the same pool as previous experiments. Three more participants (in-context condition) were replaced due to failure to complete the experiment in the time allotted.

Table 3
Experiments 4–6, Latin square orderings of motifs in each melody. Each lowercase letter refers to a single measure-long motif.

Set A		Set B	
1	`abcdefgh`	1	`ijklmnop`
2	`behfcadg`	2	`jmpnkilo`
3	`chdbgeaf`	3	`kpljomin`
4	`dfbhagce`	4	`lnjpiokm`
5	`ecgahbfd`	5	`mkoipjnl`
6	`faegbdhc`	6	`nimojlpk`
7	`gdacfheb`	7	`oliknpmj`
8	`hgfedcba`	8	`ponmlkji`

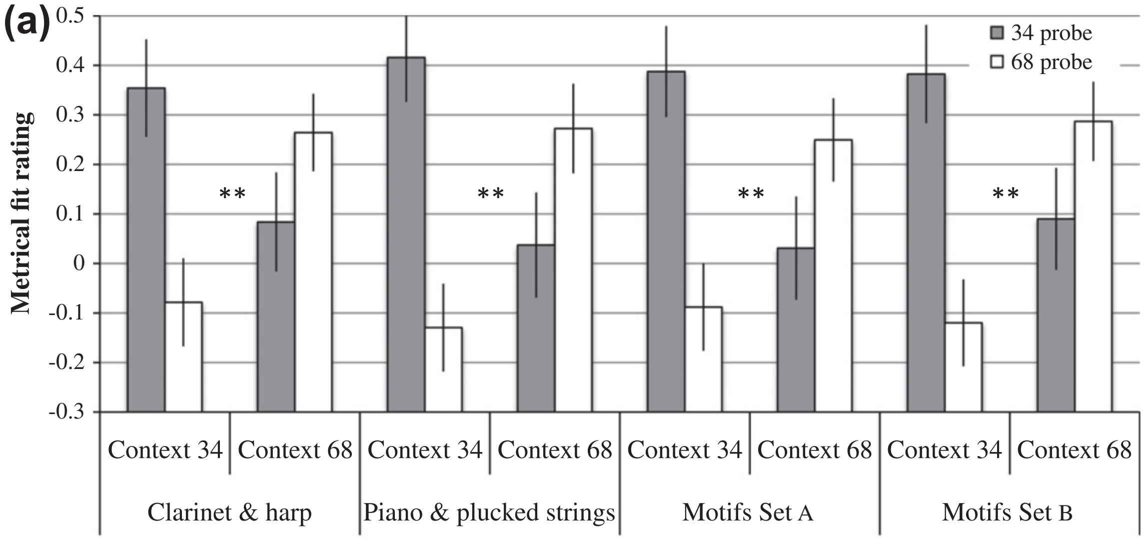
3.1.1.2. Stimuli. Melodies in Set A were composed of eight measure-long motifs, and in set B were composed of eight other motifs, using Finale software as in Part I. Motifs were arranged in a Latin square (Table 3) so that no bigrams (two motifs in a particular order) were shared across melodies. The exception to this last point was that each melody repeated once to give listeners more exposure, so that there was one extra motif conjunction. This means that a novel melody created from familiar motifs should contain at most a two-measure (two-motif) sequence that had occurred in an exposure melody. The two timbres were clarinet with harp accompaniment, and piano with plucked string accompaniment. Timbres were changed from Part I (saxophone and French horn, both wind instruments) to instruments that were more timbrally dissimilar (clarinet and piano, one a wind instrument and the other a percussive keyboard instrument). Presentation rate was set so that eighth-notes (sub-beats) were 200 ms—somewhat faster than in Part I, in an attempt to balance 3/4 and 6/8 biases.

3.1.1.3. Procedure. There was no exposure phase, only a probe phase. Half of participants ($n = 18$) heard each melody including context, followed by a metrical probe. These participants heard all melodies in all contexts (e.g. a given melody was heard with clarinet-3/4, clarinet-6/8, piano-3/4, and piano-6/8). This assessed the effectiveness of the contexts in producing a particular metrical interpretation. The remaining 18 participants heard all melodies without any context, followed by a metrical probe, just like the probe phases of following experiments. This assessed baseline preferences for each metrical probe.

3.1.2. Results

The contexts effectively elicited the desired meters, and the contextless meters produced a bias for 3/4 interpretation, much like Creel (2011, Experiment 2). For "context" listeners (Fig. 6a), an ANOVA was conducted on metrical-fit ratings with Context Meter (3/4 or 6/8), Probe Meter (3/4 or 6/8), Timbre (clarinet, piano), and Melody Set (A or B) as within-participants factors. There was a significant interaction of Context Meter × Probe Meter ($F1(1,17) = 14.35$, $p = .001$; $F2(1,14) = 95.60$, $p < .0001$; $\eta^2_G = .69$), such that listeners provided higher ratings for 3/4 probes in 3/4 contexts than in 6/8 contexts, and vice versa. The only other interaction approaching significance was Probe Instrument × Context Meter × Probe Meter ($F1(1,17) = 3.13$, $p = .095$; $F2(1,14) = 2.94$, $p = .11$; $\eta^2_G = .04$), indicating marginally greater sensitivity to the context meter for piano probes than clarinet probes. No other effects or interactions reached significance, suggesting no fundamental differences between the two timbres or motif sets.

For "no-context" listeners (Fig. 6b), an ANOVA was conducted on metrical-fit ratings with Timbre (clarinet, piano), Melody Set (A or B), and Probe Meter (3/4 or 6/8) as factors. There was an effect of Probe Meter ($F1(1,17) = 21.47$, $p = .0002$; $F2(1,14) = 648.79$, $p < .0001$; $\eta^2_G = .94$), indicating that listeners provided higher ratings overall for 3/4 metrical probes. There was also a small but significant interaction of Timbre × Melody Set × Probe Meter ($F1(1,17) = 4.48$, $p = .049$; $F2(1,14) = 4.80$, $p = .046$; $\eta^2_G = .07$). This apparently resulted from a somewhat larger 3/4 rating for clarinet-melody-set-A probes than for any of the other conditions (a rating difference of .74 relative to differences of .57, .59. and .60 in the other instrument-melody set cells). No other effects reached significance.



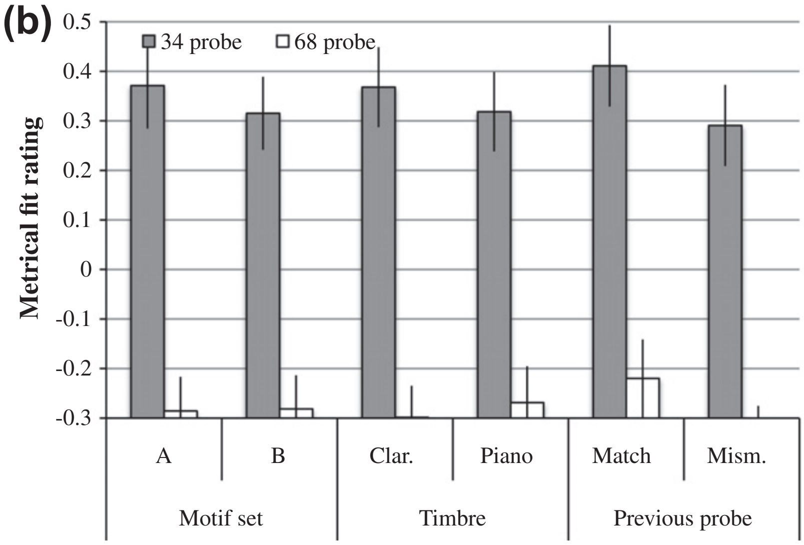


Fig. 6. Experiment 4, metrical fit ratings of melodies heard (a) in context, plotted as match–mismatch difference scores; and (b) without context, plotted as actual ratings. In (a), *y*-axis labels are in bold italic to indicate that the scale differs from earlier and later figures depicting the metrical-restoration effect—specifically, the range is almost twice the size of that in other figures (0–.5 rather than −.5 to .25, reflecting that metrical fit judgments are stronger when made immediately after hearing the meter-inducing context.

One post hoc observation is worth mentioning as it bears on the issue of context: in both sets of data, there was evidence of carry-over effects from the metrical probe on the preceding trial. For in-context listeners, if the probe meter on the previous trial matched the context meter on the current trial, ratings of the current probe were boosted (right two bars of Fig. 6a). For instance, hearing a 3/4 probe, and then a melody in 3/4 context, led to higher metrical-fit ratings for 3/4 probes than when the preceding probe was 6/8. An ANOVA on in-context ratings with Context Meter, Probe Meter, and Preceding Probe[5] (matches current context, mismatches current context) showed a significant interaction of Context Meter × Probe Meter × Preceding Probe ($F1(1,17) = 16.56$, uncorrected $p = .0008$; $F2(1,15) = 14.79$, uncorrected $p = .002$; $\eta^2_G = .22$). This interaction resulted from a stronger interaction of Context Meter × Probe Meter when the probe meter on the immediately preceding trial matched the current context meter. For no-context listeners, ratings were higher when the preceding probe matched the current one. An ANOVA with Probe Meter and Preceding Probe as factors yielded an effect

---

[5] Each participant's very first trial was omitted in this analysis.

of Preceding Probe ($F1(1,17) = 15.31$, uncorrected $p = .001$; $F2(1,15) = 15.47$, uncorrected $p = .001$; $\eta^2_G = .36$): ratings were higher when the preceding probe meter was the same as the current probe meter. These patterns suggest *sequential*-context effects in meter perception, such that the probe on the preceding trial tinged immediate perception of meter on the current trial, even if there was a metrical context present.

3.1.3. Discussion

In this experiment, listeners who were not pre-exposed to the melodies gauged the fit of metrical probes, following either melodies-in-context or melodies-out-of-context. Results suggested that the contexts elicited the intended meters, and that effects were similar for the two timbre types and for the two sets of motifs. Additionally, listeners showed a strong preference for the 3/4 probe after no-context melodies. This is consistent with results in the first three experiments, which showed some evidence of a 3/4 bias. Finally, listeners' metrical judgments seemed to be affected by the preceding metrical probe, such that that probe provided additional metrical context that modulated the following probe judgment.

Having established the suitability of the experimental materials, the next two experiments explored how timbre and motif structure in combination affect learned metrical preferences. Experiment 5 examined a case where timbre, motifs, and meter pattern together. Experiment 6 distributed motifs across meters, leaving timbre–meter patterning intact.

3.2. Experiment 5: consistent timbre–meter–motif mapping

In this experiment, the effects of consistent grouping of melodies in terms of timbres, motifs, and meter were examined. Listeners were exposed to eight melodies—four each with a characteristic timbre, motif set, and meter—and then either rated metrical probes or melody familiarity. Like Experiments 1 and 2, listeners might form timbre-specific memories of meter. Unlike Experiments 1 and 2, there was an additional component that might bolster similarity-based processing: motif content. Comparison with familiarity ratings allowed assessment of the relationship between melody recognition and metrical restoration.

3.2.1. Method

3.2.1.1. Participants. $N = 48$ participants judged metrical probes, and $N = 32$ judged familiarity. One additional participant in the metrical-judgment task was replaced because they responded identically on almost every trial, clicking at the center of the rating scale.

3.2.1.2. Stimuli. The stimuli consisted of the Part II melodies.

3.2.1.3. Procedure. Like Experiments 1–3, there was an exposure phase followed by a probe phase. In the exposure phase, each listener heard 8 of the 16 melodies. For a given listener, there were four melodies from Set A with one timbre and meter, and four from Set B with the other timbre and meter. Each of the eight melodies was presented six times, for a total of 48 exposure trials per listener. Timbre, meter, and motif set were counterbalanced across participants so that each melody occurred equally often in each cell of the design (Table 4). In the probe phase, all listeners heard all melodies, without context. For metrical-probe listeners, each melody was followed by each of the two metrical probes. For familiarity-rating listeners, melodies were not followed by metrical probes but instead a prompt to judge the familiarity of the melody. Listeners indicated familiarity ratings by clicking at a location on a 400-pixel-wide ruler, which was labeled "Definitely heard" on the left end and "Definitely didn't" on the right end. Note that this meant high familiarity corresponded to a low *x*-coordinate value, so when familiarity ratings were transformed to a -1 to +1 range, the sign was flipped such that a high value corresponded to high familiarity.

3.2.2. Results

Metrical restoration generalized across timbre for the exposed melodies, and generalized to new melodies with the original timbre–motif mappings (Fig. 7a). An ANOVA was conducted on metrical

Table 4
Experiment 5, counterbalancing of melody exposure. There were 16 lists total. Lists 9–16 were identical to Lists 1–8 except that timbres were reversed.

Set	Mel.	List 1	List 2	List 3	List 4	List 5	List 6	List 7	List 8
A melodies									
	7	Cl, 3/4	.	Cl, 6/8	.	Cl, 3/4	.	Cl, 6/8	.
	8	Cl, 3/4	.	Cl, 6/8	.	Cl, 3/4	.	Cl, 6/8	.
	1	Cl, 3/4	.	Cl, 6/8	.	.	Cl, 3/4	.	Cl, 6/8
	3	Cl, 3/4	.	Cl, 6/8	.	.	Cl, 3/4	.	Cl, 6/8
	2	.	Cl, 3/4	.	Cl, 6/8	.	Cl, 3/4	.	Cl, 6/8
	6	.	Cl, 3/4	.	Cl, 6/8	.	Cl, 3/4	.	Cl, 6/8
	4	.	Cl, 3/4	.	Cl, 6/8	Cl, 3/4	.	Cl, 6/8	.
	5	.	Cl, 3/4	.	Cl, 6/8	Cl, 3/4	.	Cl, 6/8	.
B melodies									
	7	.	Pno, 6/8	.	Pno, 3/4	Pno, 6/8	.	Pno, 3/4	.
	1	.	Pno, 6/8	.	Pno, 3/4	Pno, 6/8	.	Pno, 3/4	.
	6	.	Pno, 6/8	.	Pno, 3/4	.	Pno, 6/8	.	Pno, 3/4
	3	.	Pno, 6/8	.	Pno, 3/4	.	Pno, 6/8	.	Pno, 3/4
	2	Pno, 6/8	.	Pno, 3/4	.	.	Pno, 6/8	.	Pno, 3/4
	4	Pno, 6/8	.	Pno, 3/4	.	.	Pno, 6/8	.	Pno, 3/4
	8	Pno, 6/8	.	Pno, 3/4	.	Pno, 6/8	.	Pno, 3/4	.
	5	Pno, 6/8	.	Pno, 3/4	.	Pno, 6/8	.	Pno, 3/4	.

Note: Cl = Clarinet, Pno = piano, . = did not hear melody during exposure.

fit ratings with Learned On (clarinet = 3/4/piano = 6/8, piano = 3/4/clarinet = 6/8) as a between-participants factor, and Exposure Meter, Probe Meter, Timbre Match, and Novelty (exposed melody, new melody) as within-participants factors. There was a four-way interaction of Timbre Match × Novelty × Exposure Meter × Probe Meter ($F1(1,46) = 4.74$, $p = .03$; $F2(1,15) = 3.00$, $p = .10$; $\eta^2_G = .08$). ANOVAs on Timbre Match, Exposure Meter, and Probe Meter were conducted at each level of novelty. For familiar melodies, only the Exposure Meter × Probe Meter interaction was significant ($F1(1,47) = 12.45$, $p = .0009$; $F2(1,15) = 10.65$, $p = .005$; $\eta^2_G = .65$), suggesting that listeners inferred the exposure meter even when the timbre changed. However, for novel melodies, the Timbre Match × Exposure Meter × Probe Meter interaction was significant ($F1(1,47) = 4.98$, $p = .03$; $F2(1,15) = 11.24$, $p = .004$; $\eta^2_G = .26$). This occurred because only original-timbre trials showed the Exposure Meter × Probe Meter interaction ($F1(1,47) = 6.89$, $p = .01$; $F2(1,15) = 5.19$, $p = .038$; $\eta^2_G = .56$), while other-timbre trials did not ($F1 < 1$; $F2 < 1$). It is worth noting that the effects on probe judgments of metrical *restoration* (Fig. 7a) are smaller than the effects of actually *hearing the melodies in context* in Experiment 4. This is consistent with weaker activation of meter by the melody alone than by the full melody-in-context, implying an interaction between memory cues and surface cues.

Now familiarity ratings are considered. If familiarity predicted metrical restoration, then familiarity ratings should be high for exposed melodies regardless of timbre, but should only be high for novel melodies heard in the timbre originally paired with its motifs. However, familiarity ratings (Fig. 7b) did not bear this out: listeners showed a timbre-match effect for familiar melodies, but not for novel melodies. An ANOVA with Learned On, Exposure Meter, Timbre Match, and Novelty as within-participants factors was conducted on familiarity ratings. There was an effect of Novelty ($F1(1,30) = 80.51$, $p < .0001$; $F2(1,15) = 221.55$, $p < .0001$, $\eta^2_G = .95$), with higher familiarity ratings for exposed melodies than for novel melodies. This was qualified by an interaction of Learned On × Novelty ($F1(1,30) = 4.98$, $p = .03$; $F2(1,15) = 14.91$, $p = .002$, $\eta^2_G = .56$), which resulted from a smaller effect of novelty (still significant: $t1(15) = 7.42$, $p < .0001$; $t2(15) = 7.71$, $p < .0001$) for listeners who heard clarinet = 6/8/piano = 3/4. An effect of Timbre Match ($F1(1,30) = 15.10$, $p = .0005$; $F2(1,15) = 6.96$, $p = .019$, $\eta^2_G = .33$) was qualified by a Timbre Match × Novelty interaction ($F1(1,30) = 10.42$, $p = .003$; $F2(1,15) = 11.11$, $p = .005$, $\eta^2_G = .24$). This resulted from higher ratings of original-timbre than other-timbre for exposed melodies ($t1(31) = 5.09$, $p < .0001$; $t2(15) = 4.86$, $p = .0002$), but no difference for novel melodies ($t1(31) = .54$, $p = .59$; $t2(15) = .38$, $p = .71$). No other effects reached significance.

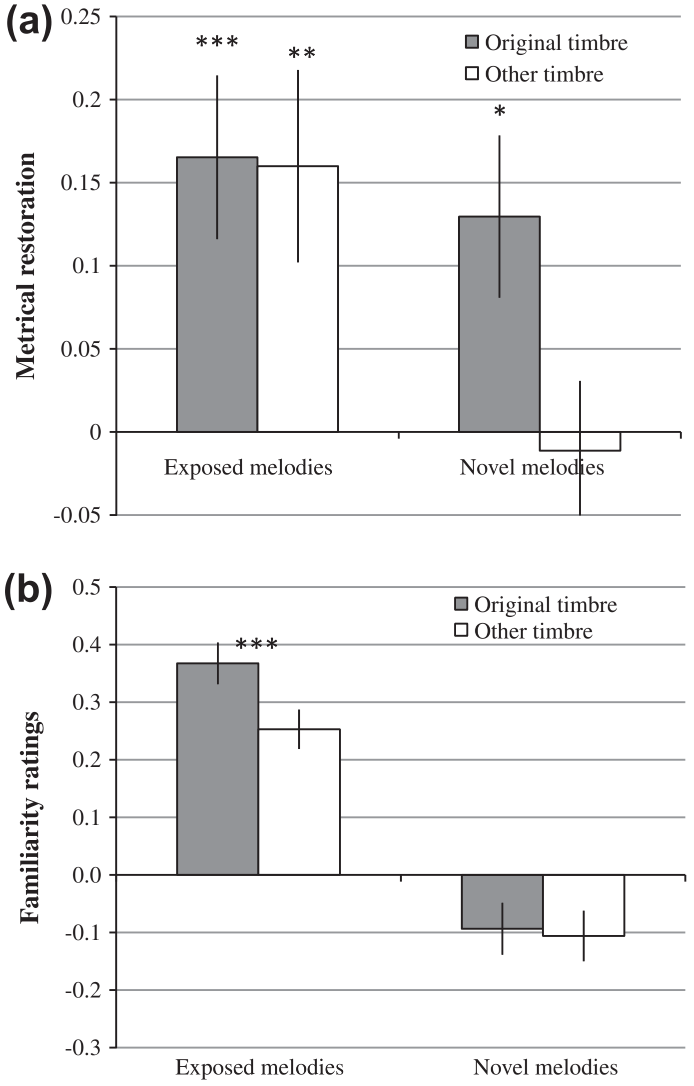


Fig. 7. Experiment 5, (a) metrical fit ratings plotted as match–mismatch difference scores, and (b) familiarity ratings.

3.2.3. Discussion

Listeners were exposed to two sets of melodies, each with a particular timbre, motif set, and meter. They restored metrical information for exposed melodies regardless of timbre. They also generalized meter preferences to new melodies, but only when melodies had *both* the original timbre and motif set of a particular meter. This is similar to the experiments in Part I in that an effect of timbre specificity was observed. However, this effect was seen only in novel melodies, which did not show metrical restoration based on timbre in Experiments 1–2. One explanation of this new pattern of generalization is that the motifs impose a tighter similarity structure than the previous melody sets in Part I—novel melodies are more similar to exposed melodies.

Interestingly, the familiarity ratings did not fully explain this pattern of results. Listeners readily distinguished familiar (exposed) melodies from unfamiliar (novel) melodies, but, unlike metrical restoration, the familiarity ratings were sensitive to timbre for the *exposed* melodies—where a timbre-specific meter effect was not seen—but not for the *novel* melodies—where a timbre-specific meter effect *was* seen. This suggests that whole-melody familiarity may not be what is driving timbre-specific metrical restoration.

One thing that seems clear is that motifs are an important element in musical similarity—which would not surprise many music theorists or music historians. However, motif structure was not the only difference between this experiment and Experiment 2: for instance, melodies were heard more often in the current experiment (six times, which was really 12 times because each melody contained two exact repetitions) than in Experiment 2 (six times without repetitions). Thus, greater amounts of exposure could be driving the effects, rather than motifs truly being a major factor in similarity structure. The next experiment eliminates motifs as a predictor of meter, to determine how important motifs are in determining metrical structure.

3.3. Experiment 6: consistent timbre, inconsistent motifs

How crucial were motifs to metrical restoration in Experiment 5? The current experiment asked what happened when motifs did *not* pattern with meter, though timbre still did. Exposure melodies were rearranged so that motifs were equally likely in either meter. It was in some ways equivalent to Experiment 2, where timbre also patterned with meter. However, in the current experiment, the similarity structure among melodies of different meters was much tighter due to shared motifs. Thus, motif content might interfere with timbre–meter learning in the current experiment, leading to weaker metrical restoration.

3.3.1. Method

3.3.1.1. Participants. $N = 48$ participants rated metrical probes, and $N = 32$ rated familiarity.

3.3.1.2. Stimuli and Procedure. These were the same as in Experiments 4–5, except that each listener heard a set of melodies in which timbre and meter were consistently paired, but each motif set was equally likely to occur with either timbre and either meter. Across participants, each melody occurred equally often in each condition (Table 5).

3.3.2. Results

Metrical probe ratings showed little evidence of metrical restoration (Fig. 8a). An ANOVA on metrical fit ratings was conducted with Exposure Meter, Probe Meter, Timbre Match, and Learned On (clarinet = 6/8/piano = 3/4; clarinet = 3/4/piano = 6/8) as factors. The Exposure Meter × Probe Meter interaction did not quite reach significance ($F1(1,46) = 2.82$, $p = .10$; $F2 < 1$; $\eta^2_G = .09$), suggesting weak if any learning. The Learned On × Probe Meter interaction was significant ($F1(1,46) = 4.57$, $p = .04$; $F2(1,15) = 37.66$, $p < .0001$; $\eta^2_G = .74$), reflecting that listeners who learned on clarinet = 6/8/piano = 3/4 gave higher ratings to 3/4 probes, while the clarinet = 3/4/piano = 6/8 other group gave higher ratings for 6/8 probes. The Timbre Match × Probe Meter interaction was also significant ($F1(1,46) = 5.78$, $p = .02$; $F2(1,15) = 6.72$, $p = .02$; $\eta^2_G = .19$), reflecting that, though there were higher

Table 5
Experiment 6, counterbalancing of melody exposure. There were 16 lists total. Lists 9–16 were identical to Lists 1–8 except that timbres were reversed.

Set	Mel.	List 1	List 2	List 3	List 4	List 5	List 6	List 7	List 8
A melodies									
	7	Cl, 3/4	.	Pno, 6/8	.	Pno, 6/8	.	Cl, 3/4	.
	8	Cl, 3/4	.	Pno, 6/8	.	Pno, 6/8	.	Cl, 3/4	.
	1	Pno, 6/8	.	Cl, 3/4	.	.	Cl, 3/4	.	Pno, 6/8
	3	Pno, 6/8	.	Cl, 3/4	.	.	Cl, 3/4	.	Pno, 6/8
	2	.	Cl, 3/4	.	Pno, 6/8	.	Pno, 6/8	.	Cl, 3/4
	6	.	Cl, 3/4	.	Pno, 6/8	.	Pno, 6/8	.	Cl, 3/4
	4	.	Pno, 6/8	.	Cl, 3/4	Cl, 3/4	.	Pno, 6/8	.
	5	.	Pno, 6/8	.	Cl, 3/4	Cl, 3/4	.	Pno, 6/8	.
B melodies									
	7	.	Pno, 6/8	.	Cl, 3/4	Cl, 3/4	.	Pno, 6/8	.
	1	.	Pno, 6/8	.	Cl, 3/4	Cl, 3/4	.	Pno, 6/8	.
	6	.	Cl, 3/4	.	Pno, 6/8	.	Pno, 6/8	.	Cl, 3/4
	3	.	Cl, 3/4	.	Pno, 6/8	.	Pno, 6/8	.	Cl, 3/4
	2	Pno, 6/8	.	Cl, 3/4	.	.	Cl, 3/4	.	Pno, 6/8
	4	Pno, 6/8	.	Cl, 3/4	.	.	Cl, 3/4	.	Pno, 6/8
	8	Cl, 3/4	.	Pno, 6/8	.	Pno, 6/8	.	Cl, 3/4	.
	5	Cl, 3/4	.	Pno, 6/8	.	Pno, 6/8	.	Cl, 3/4	.

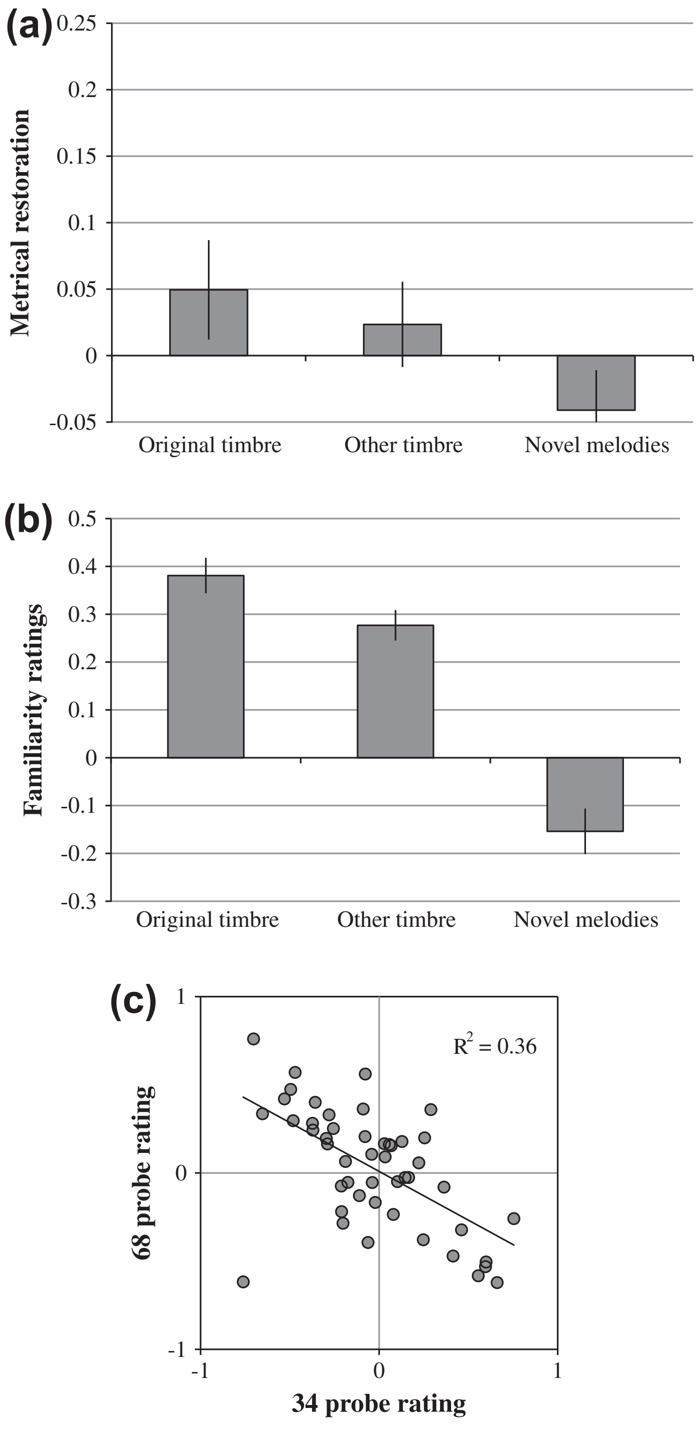


Fig. 8. Experiment 6, (a) metrical fit ratings by timbre match, plotted as match–mismatch difference scores, (b) familiarity ratings, and (c) 3/4 and 6/8 preferences by participant.

6/8 ratings for both original-timbre and other-timbre probes, the effect was much larger for original-timbre probes. It is not clear why there was a preference for 6/8 rather than 3/4, as in preceding experiments, but may be due to preexisting individual preferences for one of the two tempos, observed in previous experiments (note the wide variability in preferences in Fig. 8c).

For novel melody metrical probes, an ANOVA was conducted with Learned On, Exposure Meter, and Probe Meter. No effects approached significance in both analyses, though the Exposure Meter × Probe Meter interaction was highly significant by items ($F1(1,46) = 2.59$, $p = .12$; $F2(1,15) = 37.23$, $p < .0001$; $\eta_G^2 = .62$). However, this was due to a very small effect in the wrong direction (Fig. 8a, right).

While metrical probe ratings showed little encoding of meter, familiarity ratings (Fig. 8b) suggested excellent recognition. Planned comparisons suggested that, overall, there were higher familiarity ratings for exposed melodies than for novel melodies ($t1(31) = 9.02$, $p < .0001$; $t2(15) = 17.12$, $p < .0001$). Exposed-melody familiarity ratings were analyzed in an ANOVA with Learned On (between participants), Original Timbre (clarinet, piano), and Timbre Match (original, other) as factors. There was a main effect of Timbre Match ($F1(1,30) = 12.95$, $p = .001$; $F2(1,15) = 21.64$, $p = .0003$, $\eta^2_G = .59$), but it was qualified by a Learned On × Original Timbre × Timbre Match interaction ($F1(1,30) = 5.41$, $p = .03$; $F2(1,15) = 7.68$, $p = .01$, $\eta^2_G = .41$). This resulted from a significant Original Timbre × Timbre Match interaction in the clarinet = 3/4/piano = 6/8 condition ($F1(1,15) = 4.84$, $p = .04$; $F2(1,15) = 16.71$, $p = .0001$, $\eta^2_G = .42$), with a very strong timbre-match effect for piano melodies ($t1(15) = 4.88$, $p = .0002$; $t2(15) = 5.25$, $p < .0001$; mean difference = .25), but not for clarinet melodies ($t1(15) = .51$, $p = .61$; $t2(15) = .94$, $p = .36$; mean difference = .03). However, there was no effect of Timbre Match ($F1(1,15) = 2.02$, $p = .18$; $F2(1,15) = 7.22$, $p = .02$, $\eta^2_G = .25$; mean difference = .07) or interaction for clarinet = 6/8/piano = 3/4 melodies ($F1 < 1$; $F2 < 1$). It is not clear why there was an interaction with Learned On, but given excellent overall distinction between exposed and novel melodies, familiarity cannot explain the lack of metrical restoration.

3.3.3. Discussion

In this experiment, where motifs were distributed across meters, the metrical restoration effect was small and only marginally statistically significant. This occurred despite the pairing of *timbre* and meter—contrasting markedly with Experiment 5, which had motif consistency as well as timbre consistency, and showed robust metrical restoration. It also contrasts with Experiment 2, where timbre was consistent and there was no motif structure—in that experiment, listeners learned timbre-specific metrical information. This pattern of results suggests that the inconsistency of motif content impaired encoding or activation of metrical information, implying that motif structure is a strong determinant of metrical restoration. In terms of similarity-based activation, instances of the same motif in both meters would cause each motif to activate *both* metrical patterns, resulting in unclear metrical percepts.

Given the robust metrical-restoration effects in all preceding experiments, one might wonder whether this particular set of listeners was simply unusually inattentive. If this were the case, then listeners would presumably show a very simple response pattern (e.g., the same response repeatedly) or a thoroughly random response pattern. A quick examination of the data rules out the simple-response pattern. Another observation rules out the random-response pattern: if listeners were responding randomly, they should show no consistency in simple preferences for 3/4 vs. 6/8. However, they do: there was a strong inverse relationship between listeners' 3/4 ratings and their 6/8 ratings (Fig. 8c; $r = -.60$, $t(46) = 5.05$, $p < .0001$), suggesting consistent preferences for 3/4 vs. 6/8 for a particular participant. Additionally, there was a probe meter "bleedover" effect like that found in Experiment 4, where the previous probe interacted with judgment of the current probe (interaction of Exposure Meter × Probe Meter × Preceding Probe: $F1(1,47) = 63.26$, $p < .0001$, uncorrected; $F2(1,15) = 66.85$, $p < .0001$, uncorrected; $\eta^2_G = .60$). This bleedover effect suggests that listeners were responding contingent on their recent experience—they were paying attention. Nonetheless, they did not show metrical restoration for the tested melodies.

4. General discussion

In the Introduction, questions were raised as to whether metrical processing relies on activation of multiple similar memories rather than idealized isochronous representations, and what might count as "similar." A series of six experiments demonstrated that cross-melody similarity indeed influences metrical restoration, and that two relevant aspects of similarity are timbre and motif content. In Part I, when timbre and meter covaried—for instance, saxophone melodies always had 3/4 meter and French horn melodies always had 6/8 meters—listeners showed timbre-specific metrical restoration for familiar melodies (Experiments 1–2). However, when each timbre occurred equally often with both meters, listeners showed melody-specific but not timbre-specific metrical restoration (Experiment 3). In Part II, when both motifs and timbre were paired with meter (Experiment 5), listeners showed timbre-general metrical restoration for familiar melodies, and timbre-specific generalization to novel melodies containing the motifs. However, when timbre was paired with meter and motifs varied freely (Experiment 6), metrical learning was not observed, despite attentive participants and excellent recognition accuracy. Simple *recognition* of melodies (Experiments 5–6) did not predict patterns of metrical restoration.

This series of experiments shows that the similarity structure within a listener's musical memories can shape metrical processing and musical recognition. Consistent with previous research (Halpern & Müllensiefen, 2007; Peretz et al., 1998; Poulin-Charronnat et al., 2011), familiarity judgments showed effects of timbre. The more novel results are that multiple musical memories seem to shape processing depending on the timbre and motif characteristics of listeners' exposure, and that simple recognition does not predict metrical restoration. This result expands Creel's (2011) finding of melody-specific metrical restoration to suggest a rich, similarity-based representational system (e.g. Hintzman, 1986; Pearce & Wiggins, 2004, 2006) which actively shapes metrical processing, and possibly processing of timing generally. This type of memory organization could support processing and generation of differential expectations in diverse musical styles.

The finding that timbre shapes musical memory activation is intriguing, since timbre is sometimes construed as less central to musical representations than pitch or meter. This is similar to the role talker information in word recognition (Creel et al., 2008; Goldinger 1996, 1998; Palmeri et al., 1993) or color information in object recognition (Goffaux et al., 2005; Mitterer & De Ruiter, 2008; Oliva & Schyns, 2000; Tanaka & Presnell, 1999). The current study suggests that the importance of timbre to metrical information is determined by the structure of the listener's musical similarity space. Shared motif content contributed to similarity as well, in that motifs seemed to allow metrical restoration to generalize to new melodies, and also blocked metrical restoration when motifs occurred in both meters. This is consistent with each motif activating whichever meter(s) it is associated with. If it is associated with one meter, that meter is activated strongly, but if it is associated with two meters, they must compete for activation, with neither dominating perception.

This study also suggests a non-obvious relationship between familiarity and metrical processing. Good (or poor) melody recognition does not cleanly predict whether metrical restoration will (or will not) occur. Familiarity seems to reflect properties of whole melodies, while metrical restoration seems to reflect similarity-based activation of melody subcomponents. For instance, familiarity ratings are influenced by timbre similarity. However, metrical activation seems to operate by a somewhat-different set of rules: it can happen when whole-melody familiarity is low (Experiment 5), and it can be absent when whole-melody familiarity is high (Experiment 6). Without making strong claims about nature of the processing going on, it is interesting to note that this dissociation mirrors Church and Schacter's (1994; Schacter & Church, 1992) implicit/explicit dissociation of talker specificity on word processing.[6] They found that performance on implicit tasks—stem completion (Schacter and Church) and identification of low-pass-filtered words (Church &and Schacter)—was aided by same-talker presentation, while overt recognition was not. Here, motif similarity influenced metrical restoration but not familiarity. This raises the possibility that metrical restoration is calculated on a more implicit basis than recognition. Proving this definitively would require different tasks than the ones used here, such as showing different responses of familiarity judgments and meter judgments to a levels-of-processing manipulation, suggesting interesting directions for future work.

4.1. Implications for learning about temporal events

One interpretation of these results is that periodicity is not calculated based solely on the external signal; it is calculated based on the signal itself in combination with activated memories. This means that listeners can fill in periodic activity from an ambiguous stimulus—a melody heard out of context, a spoken sentence that is not truly isochronous (see Pitt and Samuel (1990) for a demonstration of this), even an ambulating predator (or prey) partially obscured by foliage. Rather than assuming some cognitive predisposition to detect periodicity, some types of periodicity detection might be more succinctly ascribed to memory activation. Memories, in turn, may tend toward periodicity due to the pre-

---

[6] Thanks to Arty Samuel for pointing out this connection.

ponderance of (quasi-)periodic environmental events (see Schwartz, Howe, and Purves (2003) for a similar argument that perception of musical consonance might stem from exposure to periodic stimuli [speech] in the environment).

Ascribing periodicity to memory rather than (or in addition) to some intrinsic property of the mind/brain has the potential to explain periodic perceptual phenomena that are not currently understood. For instance, why do listeners perceive spoken language to be isochronous when it demonstrably is not (e.g. Cummins & Port, 1998)? It may be that the sum of activated memories of spoken language generate a metrical "echo" (Hintzman, 1986) or resonance (Carpenter, Grossberg, & Rosen, 1991) in recognition, shaping the listener's phenomenal experience. As noted in the introduction, there is not a good periodicity explanation of non-isochronous or "complex" meters, which rarely occur in Western music but are common in other musical cultures (West African, Indian, Eastern European; see London (1995) for discussion).[7] A memory-based account of meter, rather than an account focusing exclusively on entrainment to isochrony, may be a more effective explanation for perception of complex (non-isochronous) meters. That is, when hearing music in a non-isochronous meter, memories of similar non-isochronous melodies or melodic chunks are activated, supporting processing of the meter.

Finally, similarity-based metrical restoration may contribute to phenomena like the illusory perception of amplitude or length differences in repeating auditory patterns (Brochard et al., 2003; Iversen et al., 2009). Listeners may simply be coactivating memory traces that "fill in" a metrical percept, either musical traces which fill in a strong–weak pattern (Brochard et al.), or, perhaps, spoken-language traces which fill in an iambic pattern for listeners in one language vs. a trochaic pattern for listeners in another language (Iversen et al., 2008). Put more simply, perhaps what the brain is doing in a variety of metrical entrainment studies is perceptual restoration: it is coactivating or simulating the original conditions in which similar events have been experienced.

4.2. Limitations

The account given in this study is that metrical restoration effects are sensitive to coherence among multiple memory representations. However, the reader should keep in mind a few limitations or alterative explanations of these patterns of results. The most powerful alternative hypothesis is that listeners are detecting patterns in the exposure phase, and are responding based on strategy rather than implicit memory activation. For instance, in Experiment 1, listeners might notice that the saxophone melodies are all in 3/4 and French horn melodies are all in 6/8. The listener might thus give saxophone melodies high 3/4 ratings and low 6/8 ratings at test (and the reverse for French horn melodies). Though this cannot be definitively disproven, it seems quite unlikely. First, listeners were not alerted to attend to any aspect of the music in particular until after the end of the exposure phase. Second, if listeners did use a strategy based on timbre–meter patterning, they would presumably have applied the strategy to both old *and* new melodies with the same timbre, which they did not (Experiments 1–2). Third, if listeners used a timbre–meter patterning strategy, they presumably would not be misled by ambiguous motif structure, which they were (Experiment 6). Thus, it seems unlikely that listeners' metrical judgments result from strategic processing.

A second limitation is that this study does not distinguish between rhythm and meter. That is, listeners could be storing the rhythms of the accompanying contexts—the exact durations of succeeding notes—rather than truly metrical information, the repeating pattern of strong and weak beats.[8] This seems plausible, and such a result would be consistent with the construal of meter suggested here: that

---

[7] Note that these meters *do* possess isochrony at the sub-beat level, but not at the level of the tactus, where one would clap or step in time.

[8] Yet another possibility, raised by a reviewer, is that listeners are simply counting number of events per bar, and matching memories of higher note counts to the 3/4 metrical probes (three events per bar) and lower counts to 6/8 metrical probes (2 events per bar). To check whether this was a possible counter explanation, onsets were counted for all stimuli. In all cases, note counts were nearly the same in 3/4 vs. 6/8 contexts, and all of the very small differences were in the wrong direction. For instance, for Part 1 stimuli, 3/4 contexts had on average 4.03 onsets per bar (SD = 1.06), while 6/8 had on average 4.17 (SD = 1.31) onsets per bar. The pattern was identical regardless of whether the melody and context were counted together or separately, both for Part I and Part II stimuli. Thus, metrical restoration results could not have occurred based on listeners matching *smaller* note counts to 6/8 metrical probes.

meter itself is an emergent property, a composite, of numerous rhythmic experiences. It would be interesting to know whether consistent but *non*-isochronous rhythmic contexts presented as probes would receive higher probe ratings than metrical probes. This would be consistent with the meter-as-emergent-rhythmic-property account.

Relatedly, this study does not distinguish between *tempo*—the absolute inter-beat interval—and *meter*—how the tactus lines up with the subbeats of the melody. That is, the two meters in each experiment were also two different rates. This would be seemingly easy to test: if, say, a melody previously presented in a 6/8 context were sped up to the inter-beat interval of a 3/4 melody, and then a 6/8 probe is given, that probe will be a better tempo match for the 3/4 melody, but a better meter match for the 6/8 melody. My suspicion is that the results would actually be somewhat complex, reflecting a combination of metrical match, preexisting preferences, and degree of match to the original tempo (presentation rate). While this is an interesting task for future exploration, is beyond the scope of the present study.

A third limitation is that an alternative associative explanation is not ruled out—that listeners are simply associating correlated elements in memory, rather than specifically encoding metrical information. For instance, listeners might verbally label each melody they hear (explicitly or implicitly) as "saxophone, fast beat rate" or "French horn, slow beat rate." They then make probe matches based on these verbally-encoded cues, though such cues must be bolstered by recognition of the melody itself as familiar (given that they do not make timbre-based metrical probe matches to unfamiliar melodies in Experiments 1–2). This would predict that an exposure phase that gave only verbal information about timbre and meter might generate the same test results, or that less-labelable stimuli would show weaker effects. While this explanation cannot be ruled out completely based on the current data, it seems unlikely, given that simple meter-timbre associations ("saxophone = 3/4") are ruled out by lack of results on novel melodies in Experiments 1–2, and given that timbre–meter associations appeared *not* to be effective when motifs conflicted (Experiment 6). Relatedly, the study cannot demonstrate for certain that listeners truly *experience the meters* as being 3/4-like or 6/8-like depending on filled-in context. This would be more effectively addressed with an overt metrical production task, or a brain-response paradigm such as that used by Brochard et al. (2003).

Finally, a remaining question is how memory specificity effects in metrical perception might change over time. Do timbre-specific memory representations persist for the long term, or do they disappear after a few hours or after sleep consolidation? Do motif-specific representations persist for longer than timbre-specific ones? Further, does the metrical restoration effect at all endure after sleep consolidation, or do memories become so strongly integrated with listeners' existing musical knowledge—akin to Gaskell and Dumay (2003) findings of integration of word representations into the lexicon—that the effect of specific exposure weakens or disappears? These are interesting questions for future research.

5. Conclusion

The current study demonstrates that meter perception, previously thought to be determined from idealized underlying representations, is influenced by multiple memory representations. Both timbre and motif content shape metrical restoration, and metrical restoration is not directly tied to melody familiarity. These results together suggest that musical memory is complex, acoustically-detailed, and highly interactive, consistent with similarity-based models of recognition (Goldinger, 1998; Hintzman, 1986; Pearce & Wiggins, 2004, 2006). This adds to the growing, if controversial, literature suggesting that many properties of perceptual objects such as color in vision (Goffaux et al., 2005; Oliva & Schyns, 2000) or talker information in speech (Creel et al., 2008; Goldinger, 1996, 1998; Palmeri et al., 1993) can facilitate recognition. Further, this work suggests that previous demonstrations of highly-specific memory representations are not just isolated impressive feats of memory, but serve the function of facilitating context-specific processing.

Acknowledgments

This work was supported in part by a Hellman Foundation Fellowship to S.C.C. Thanks to Jeff Elman for helpful discussion.

Appendix A

Music background of participants and correlations of music experience with observed experimental effects.

Experiment	Data available		Years playing an instrument			Took courses[a]	*r*[b]	*p*
	%	*n/N*	*M*	*SD*	Range			
1	86	31/36	5.3	4.4	0–16	5	−.06 (fam. timb.)	.77
2	72	26/36	5.8	4.4	0–14	3	.07 (fam. timb.)	.75
3	86	31/36	3.3	4.2	0–16	5	.24	.18
4, Context	72	13/18	4.9	5.3	0–14	3	.13	.67
4, No context (3/4 pref.)	50	9/18	5.0	5.9	0–13	3	.87	**.002**
5, Meter	85	41/48	4.9	4.3	0–14	5	.37 (fam. mels.)[c] .20 (nov. mels.)	**.02** .21
5, Recog.	84	27/32	6.5	5.1	0–17	8	.16	.42
6, Meter	85	41/48	3.5	4.0	0–12	5	.14	.17
6, Recog.	88	28/32	4.3	5.1	0–18	3	.52[d]	**.005**

Note. Fam. = familiar, timb. = timbre, pref. = preference, mels. = melodies, recog. = recognition.

[a] Courses in music theory or music composition.

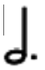
[b] Correlation between years of music experience and metrical restoration effect magnitude, or between years experience and familiar–unfamiliar ratings.

[c] This seems to be carried by two very high music experience participants, who produced metrical restoration scores more than two SDs above other participants.

[d] Despite this correlation, 27/28 participants gave higher familiar ratings to previously-heard melodies than to unheard melodies, suggesting that even musical novices recognized melodies above chance. The one participant who showed a familiarity-rating difference two SDs above the mean also had the most music experience (18 years of playing an instrument).

Appendix B

Rhythmic patterns used in Part I (Experiments 1–3).

Musical notation	Note durations (subbeats)
	6
	4 1 1
	3 1 2
	3 1 1 1
	2 1 1 2
	2 1 1 1 1
	1 1 1 1 2
	1 1 1 1 1 1

Note. While there were just 8 rhythms, if contour is taken into account, there were 54 different rhythm + contour combinations (like the 16 motifs in Part II).

Appendix C

Table C1
Metrical fit ratings by condition in Experiment 1.

Melodies	Timbre	Probe presented			
		34 Probe		68 Probe	
		Mean	*SD*	Mean	*SD*
Exposure melodies	Original timbre				
	Expose 34	0.214	*0.294*	–0.017	*0.296*
	Expose 68	0.063	*0.306*	0.073	*0.351*
	Other timbre				
	Expose 34	0.107	*0.369*	0.007	*0.297*
	Expose 68	0.051	*0.290*	0.042	*0.311*
	Novel timbre				
	Expose 34	0.202	*0.312*	–0.007	*0.298*
	Expose 68	0.139	*0.296*	0.017	*0.327*
Novel melodies	Timbre 34	0.147	*0.327*	0.007	*0.306*
	Timbre 68	0.121	*0.313*	–0.024	*0.277*
	Novel timbre	0.173	*0.345*	–0.084	*0.333*

Table C2
Metrical fit ratings by condition in Experiment 2.

Melodies	Timbre	Probe presented			
		34 Probe		68 Probe	
		Mean	*SD*	Mean	*SD*
Exposure melodies	Original timbre				
	Expose 34	0.240	*0.356*	–0.056	*0.360*
	Expose 68	0.073	*0.364*	0.127	*0.359*
	Other timbre				
	Expose 34	0.167	*0.325*	0.072	*0.402*
	Expose 68	0.153	*0.350*	0.079	*0.278*
Novel melodies	Timbre 34	0.145	0.145	*0.362*	–0.005
	Timbre 68	0.106	*0.320*	0.044	*0.332*

Table C3
Metrical fit ratings by condition in Experiment 3.

Melodies	Timbre	Probe presented			
		34 Probe		68 Probe	
		Mean	*SD*	Mean	*SD*
Exposure melodies	Original timbre				
	Expose 34	0.205	*0.317*	0.115	*0.323*
	Expose 68	–0.066	*0.316*	0.051	*0.272*
	Other timbre				
	Expose 34	0.237	*0.298*	0.111	*0.306*
	Expose 68	–0.053	*0.295*	0.003	*0.324*
Novel melodies		0.090	*0.349*	–0.019	*0.311*

Table C4
Metrical fit ratings by condition in Experiment 4.

Melodies	Timbre	Probe presented			
		34 Probe		68 Probe	
		Mean	*SD*	Mean	*SD*
Clarinet and harp	Context 34	0.354	*0.418*	−0.078	*0.377*
	Context 68	0.084	*0.425*	0.264	*0.332*
	No context	0.368	*0.343*	−0.299	*0.271*
Piano and plucked strings	Context 34	0.416	*0.380*	−0.129	*0.376*
	Context 68	0.037	*0.451*	0.272	*0.384*
	No context	0.318	*0.340*	−0.269	*0.312*
Motifs Set A	Context 34	0.387	*0.390*	−0.088	*0.374*
	Context 68	0.031	*0.443*	0.249	*0.357*
	No context	0.371	*0.369*	−0.286	*0.293*
Motifs Set B	Context 34	0.382	*0.421*	−0.120	*0.372*
	Context 68	0.090	*0.437*	0.287	*0.340*
	No context	0.315	*0.313*	−0.281	*0.288*
Previous probe 34	Context 34	0.436	*0.376*	−0.121	*0.388*
	Context 68	0.124	*0.440*	0.218	*0.371*
	No context	0.411	*0.349*	−0.338	*0.267*
Previous probe 68	Context 34	0.340	*0.430*	−0.090	*0.378*
	Context 68	0.009	*0.459*	0.320	*0.336*
	No context	0.291	*0.348*	−0.220	*0.333*

Table C5
Metrical fit ratings by condition in Experiment 5.

Melodies	Timbre	Probe presented			
		34 Probe		68 Probe	
		Mean	*SD*	Mean	*SD*
Exposure melodies	Original timbre				
	Expose 34	0.216	0.424	−0.084	0.350
	Expose 68	0.013	0.382	0.043	0.377
	Other timbre				
	Expose 34	0.197	0.425	−0.123	0.382
	Expose 68	0.043	0.473	0.042	0.412
Novel melodies	Original timbre				
	Motifs 34	0.169	0.384	−0.072	0.380
	Motifs 68	0.046	0.415	0.066	0.393
	Other timbre				
	Motifs 34	0.083	0.382	−0.030	0.386
	Motifs 68	0.079	0.442	−0.057	0.374

Table C6
Metrical fit ratings by condition in Experiment 6.

Melodies	Timbre	Probe presented			
		34 Probe		68 Probe	
		Mean	*SD*	Mean	*SD*
Exposure melodies	Original timbre				
	Expose 34	−0.077	*0.416*	0.000	*0.399*
	Expose 68	−0.102	*0.441*	0.075	*0.386*
	Other timbre				
	Expose 34	0.007	*0.415*	−0.003	*0.378*
	Expose 68	0.007	*0.438*	0.043	*0.376*
Novel melodies	Timbre 34				
	Timbre 68	−0.046	*0.383*	0.042	*0.377*
		−0.003	*0.400*	0.005	*0.358*

Appendix D

Motifs used in Part II (Experiments 4–6).

Set A			Set B		
Notation[a]	Durations[b]	Contours[c]	Notation[a]	Durations[b]	Contours[c]
	2 1 1 2	UDD		4 1 1	DU
	2 1 1 1 1	DDUU		1 1 1 1 1 1	UUUDD
	2 1 1 2	DUU		4 1 1	UD
	3 1 2	DU		3 1 2	UD
	3 1 1 1	DUD		3 1 2	UU
	2 1 1 1 1	UUDD		3 1 1 1	DUU
	2 1 1 2	UUU		1 1 1 1 1 1	DUDUD
	3 1 1 1	UDD		1 1 1 1 1 1	DDDUU
Total	33	U = 13, D = 12	Total	34	U = 14, D = 12

---

[a] Melodic contours are only specified at the note-to-note level—pitch direction between note n and note $n + 2$, for example, was free to vary.

[b] Durations are in sub-beat (200 ms) units.

[c] Contours indicate pitch direction between adjacent notes; U = up, D = down.

References

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*(3), 379–384.

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology, 20*(1), 38–64. <http://www.ncbi.nlm.nih.gov/pubmed/3338267>.

Boersma, P., & Weenink, D. (2007). Praat: Doing phonetics by computer (Version 5.1.44) [Computer program]. <http://www.praat.org/> Retrieved 04.10.10.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences, 105*(38), 14325–14329. http://dx.doi.org/10.1073/pnas.0803390105.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science, 14*(4), 362–366. <http://www.ncbi.nlm.nih.gov/pubmed/12807411>.

Carpenter, G. A., Grossberg, S., & Rosen, D. B. (1991). ART 2-A: An adaptive resonance algorithm for rapid category learning and recognition. *Neural Networks, 4*, 493–504.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General, 113*(3), 394–412.

Church, B. A., & Schacter, D. L. (1994). Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(3), 521–533. <http://www.ncbi.nlm.nih.gov/pubmed/8207370>.

Clopper, C. G., & Pisoni, D. B. (2004). Effects of talker variability on perceptual learning of dialects. *Language and Speech, 47*, 207–239.

Creel, S. C. (2011). Specific previous experience affects perception of harmony and meter. *Journal of Experimental Psychology: Human Perception and Performance*. http://dx.doi.org/10.1037/a0023507.

Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2008). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition, 106*, 633–664. http://dx.doi.org/10.1016/j.cognition.2007.03.013.

Creel, S. C., & Tumlin, M. A. (2011). On-line acoustic and semantic interpretation of talker information. *Journal of Memory and Language*. http://dx.doi.org/10.1016/j.jml.2011.06.005.

Cummins, F., & Port, R. F. (1998). Rhythmic constraints on stress timing in English. *Journal of Phonetics, 26*, 145–171.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language, 31*(2), 218–236. http://dx.doi.org/10.1016/0749-596X(92)90012-M.

Deutsch, D. (1972). Octave generalization and tune recognition. *Perception & Psychophysics, 11*, 411–412.

DeWitt, L. A., & Samuel, A. G. (1990). The role of knowledge-based expectations in music perception: Evidence from musical restoration. *Journal of Experimental Psychology: General, 119*(2), 123–144. <http://www.ncbi.nlm.nih.gov/pubmed/2141351>.

Dilley, L. C., & McAuley, J. D. (2008). Distal prosodic context affects word segmentation and lexical processing. *Journal of Memory and Language, 59*(3), 294–311. http://dx.doi.org/10.1016/j.jml.2008.06.006.

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics, 67*(2), 224–238.

Ellis, R. J., & Jones, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 264–280. http://dx.doi.org/10.1037/a0013482.

Elman, J. L., & McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language, 165*, 143–165.

Essens, P. J., & Povel, D.-J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception & Psychophysics, 37*(1), 1–7. <http://www.ncbi.nlm.nih.gov/pubmed/3991313>.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89*, 105–132.

Gjerdingen, R., & Perrott, D. (2008). Scanning the dial: The rapid recognition of music genres. *Journal of New Music Research, 37*(2), 93–100. http://dx.doi.org/10.1080/09298210802479268.

Goffaux, V., Jacques, C., Mouraux, A., Oliva, A., Schyns, P., & Rossion, B. (2005). Diagnostic colours contribute to the early stages of scene categorization: Behavioural and neurophysiological evidence. *Visual Cognition, 12*(6), 878–892. http://dx.doi.org/10.1080/13506280444000562.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*(5), 1166–1183.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*(2), 251–279.

Halpern, A. R., & Müllensiefen, D. (2007). Effects of timbre and tempo change on memory for music. *The Quarterly Journal of Experimental Psychology, 61*(9), 1371–1384. http://dx.doi.org/10.1080/17470210701508038.

Handel, S., & Lawson, G. R. (1983). The contextual nature of rhythmic interpretation. *Perception & Psychophysics, 34*(2), 103–120. <http://www.ncbi.nlm.nih.gov/pubmed/6634366>.

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 30*(5), 956–974. http://dx.doi.org/10.1037/0096-1523.30.5.956.

Hannon, E. E., & Trehub, S. E. (2005a). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences, 102*(35), 12639–12643.

Hannon, E. E., & Trehub, S. E. (2005b). Metrical categories in infancy and adulthood. *Psychological Science, 16*(1), 48–55.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York: Wiley.

Hintzman, D. L. (1986). Schema abstraction in a multiple-trace memory model. *Psychological Review, 93*, 411–428.

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America, 124*(4), 2263–2271. http://dx.doi.org/10.1121/1.2973189.

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences, 1169*, 58–73. http://dx.doi.org/10.1111/j.1749-6632.2009.04579.x.

Iverson, P., & Krumhansl, C. L. (1993). Isolating the dynamic attributes of musical timbre. *Journal of the Acoustical Society of America, 94*(5), 2595–2603. <http://www.ncbi.nlm.nih.gov/pubmed/8270737>.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*(5), 323–355.

Jones, M. R., Moynihan, H., Mackenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13*, 313–319.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology, 51*(2), 141–178. http://dx.doi.org/10.1016/j.cogpsych.2005.05.001.

Krumhansl, C. L. (2010). Plink: "Thin slices" of music. *Music Perception, 27*(5), 337–354.

Krumhansl, C. L., Louhivuori, J., Toiviainen, P., Järvinen, T., & Eerola, T. (1999). Melodic expectation in Finnish spiritual folk hymns: Convergence of statistical, behavioral, and computational approaches. *Music Perception, 17*(2), 151–195.

Krumhansl, C. L., Toivanen, P., Eerola, T., Toiviainen, P., Järvinen, T., & Louhivuori, J. (2000). Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition, 76*(1), 13–58. <http://www.ncbi.nlm.nih.gov/pubmed/10822042>.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119–159.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.

Levitin, D. J. (1994). Absolute memory for musical pitch: Evidence from the production of learned melodies. *Perception & Psychophysics, 56*(4), 414–423.

Levitin, D. J. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics, 58*(6), 927–935.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/ – II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America, 94*(3 Pt 1), 1242–1255. <http://www.ncbi.nlm.nih.gov/pubmed/8408964>.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /1/: A first report for publication. *Journal of the Acoustical Society of America, 89*(2), 874–886.

London, J. (1995). Some examples of complex meters and their implications for models of metric perception. *Music Perception, 13*(1), 59–77.

McAuley, D. (2011). Movement and the auditory encoding of rhythm. *Abstract of paper presented at the 10th annual auditory perception, cognition, and action meeting*, Seattle, WA.

McAuley, J. D., Henry, M. J, Rajarajan, P., & Nave, K. (2011). Effect of movement on the metrical interpretation of ambiguous rhythms: Phillips-Silver and Trainor (2007) revisited. In *Presented at the biennial meeting of the society for music perception and cognition*. August 11-14. Rochester, New York: Eastman School of Music.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*(1), 1–86. <http://www.ncbi.nlm.nih.gov/pubmed/3753912>.

McLennan, C. T., & Luce, P. A. (2005). Examining the time course of indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(2), 306–321. http://dx.doi.org/10.1037/0278-7393.31.2.306.

McLennan, C. T., Luce, P. A., & Charles-Luce, J. (2003). Representation of lexical form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 539–553. http://dx.doi.org/10.1037/0278-7393.29.4.539.

Mitterer, H., & De Ruiter, J. P. (2008). Recalibrating color categories using world knowledge. *Psychological Science, 19*(7), 629–634. http://dx.doi.org/10.1111/j.1467-9280.2008.02133.x.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods, 8*(4), 434–447. http://dx.doi.org/10.1037/1082-989X.8.4.434.

Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology, 41*(2), 176–210. http://dx.doi.org/10.1006/cogp. 1999.0728.

Palmer, C., Jungers, M. K., & Jusczyk, P. W. (2001). Episodic memory for musical prosody. *Journal of Memory and Language, 45*(4), 526–545. http://dx.doi.org/10.1006/jmla.2000.2780.

Palmeri, T. J., Goldinger, S. D., & Pisoni, David. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(2), 309–328.

Patel, A. D. (2008). *Music, language, and the brain*. NY: Oxford University Press.

Pearce, M., & Wiggins, G. A. (2004). Improved methods for statistical modelling of monophonic music. *Journal of New Music Research, 33*(4), 367–385. http://dx.doi.org/10.1080/0929821052000343840.

Pearce, M. T., & Wiggins, G. A. (2006). Expectation in melody: The influence of context and learning. *Music Perception, 23*(5), 377–405.

Peretz, I., Gaudreau, D., & Bonnel, A. M. (1998). Exposure effects on music preference and recognition. *Memory & Cognition, 26*(5), 884–902. <http://www.ncbi.nlm.nih.gov/pubmed/9796224>.

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science, 308*(5727), 1430. http://dx.doi.org/10.1126/science.1110922.

Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition, 105*(3), 533–546. http://dx.doi.org/10.1016/j.cognition.2006.11.006.

Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition, 67*(1), 94–102. http://dx.doi.org/10.1016/j.bandc.2007.11.007.

Pitt, M. A., & Samuel, A. G. (1990). The use of rhythm in attending to speech. *Journal of Experimental Psychology: Human Perception and Performance, 16*(3), 564–573. <http://www.ncbi.nlm.nih.gov/pubmed/2144571>.

Poulin-Charronnat, B., Bigand, E., Lalitte, P., Madurell, F., Vieillard, S., & McAdams, S. (2011). Effects of a change in instrumentation on the recognition of musical materials. *Music Perception, 22*(2), 239–263.

Povel, D.-J., & Essens, P. J. (1985). Perception of temporal patterns. *Music Perception, 2*(4), 411–440.

Randel, D. M. (1986). *The new Harvard dictionary of music*. Cambridge, MA: Harvard UP.

Remez, R. E., Rubin, P. E., Pisoni, D. B., & Carrell, T. D. (1981). Speech perception without traditional speech cues. *Science, 212*, 947–949.

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology: General, 110*(4), 474–494. <http://www.ncbi.nlm.nih.gov/pubmed/2939193>.

Saygin, A. P., Driver, J., & de Sa, V. R. (2008). In the footsteps of biological motion and multisensory perception: Judgments of audiovisual temporal relations are enhanced for upright walkers. *Psychological Science, 19*(5), 469–475. http://dx.doi.org/10.1111/j.1467-9280.2008.02111.x.

Schacter, D. L., & Church, B. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 915–930. <http://www.ncbi.nlm.nih.gov/pubmed/1402716>.

Schellenberg, E. G., & Trehub, S. E. (2003). Good pitch memory is widespread. *Psychological Science, 14*(3), 262–266. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12741751..

Schellenberg, E. G., Iverson, P., & McKinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin & Review, 6*(4), 641–646. <http://www.ncbi.nlm.nih.gov/pubmed/10682207>.

Schwartz, D. A., Howe, C. Q., & Purves, D. (2003). The statistical structure of human speech sounds predicts musical universals. *Journal of Neuroscience, 23*(18), 7160–7168.

Smith, N. A., & Schmuckler, M. A. (2008). Dial A440 for absolute pitch: Absolute pitch memory by non-absolute pitch possessors. *The Journal of the Acoustical Society of America, 123*(4), EL77–84. http://dx.doi.org/10.1121/1.2896106.

Snyder, J. S., & Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cognitive Brain Research, 24*(1), 117–126. http://dx.doi.org/10.1016/j.cogbrainres.2004.12.014.

Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology, 25*(2), 207–222.

Sumner, M., & Samuel, A. G. (2005). Perception and representation of regular variation: The case of final /t/. *Journal of Memory and Language, 52*(3), 322–338. http://dx.doi.org/10.1016/j.jml.2004.11.004.

Tanaka, J. W., & Presnell, L. M. (1999). Color diagnosticity in object recognition. *Perception & Psychophysics, 61*(6), 1140–1153. <http://www.ncbi.nlm.nih.gov/pubmed/10497433>.

Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science, 167*(3917), 392–393.

Yoshida, K. A., Iversen, J. R., Patel, A. D., Mazuka, R., Nito, H., Gervain, J., et al (2010). The development of perceptual grouping biases in infancy: A Japanese–English cross-linguistic study. *Cognition, 115*(2), 356–361. http://dx.doi.org/10.1016/j.cognition.2010.01.005.