UC San Diego
UC San Diego Previously Published Works

Title
Metrical Restoration From Local and Global Melodic Cues

Permalink

Journal

ISSN

Author
Creel, Sarah C

Publication Date
2020-11-25

DOI

Peer reviewed

Metrical restoration from local and global melodic cues:

Rhythmic patterns and overall melodic form

Sarah C. Creel

University of California San Diego

Department of Cognitive Science

9500 Gilman Drive Mail Code 0515

La Jolla CA 92093-0515

Phone: 858-534-7308 Fax: 858-534-1128

Email: screel@ucsd.edu

Abstract

What factors influence listeners' perception of meter in a musical piece or a musical style? Many cues are available in the musical "surface," i.e., the pattern of sounds physically present during listening. Models of meter processing focus on the musical surface. However, percepts of meter and other musical features may also be shaped by reactivation of previously-heard music, consistent with exemplar accounts of memory. The current study explores a phenomenon that is here termed *metrical restoration*: listeners who hear melodies with ambiguous meters report meter preferences that match previous listening experiences in the lab, suggesting *reactivation* of those experiences. Previous studies suggested that timbre and brief rhythmic patterns may influence metrical restoration. However, variations in the magnitude of effects in different experiments suggest that other factors are at work. Experiments reported here explore variation in metrical restoration as a function of: *melodic diversity* in timbre and tempo; associations of *rhythmic patterns* with particular melodies and meters; and associations of meter with *overall melodic form*. Rhythmic patterns and overall melodic form, but not timbre, had strong influences. Results are discussed with respect to style-specific or culture-specific musical processing, and everyday listening experiences. Implications for models of musical memory are also discussed.

When listening to music, its pattern of strong and weak beats—its meter—seems to emerge from the music itself. However, imperfect attempts at machine recognition of beats (see, e.g., Casey, Veltkamp, Goto, Leman, Rhodes, & Slaney, 2008; Scheirer, 1998) and poor performance in processing metrical patterns that do not occur in one's culture (Hannon & Trehub, 2005a, 2005b; Hannon, Soley, & Ullal, 2012; Kalender, Trehub, & Schellenberg, 2013) suggest that musical surface cues are not sufficient. Prior experience with musical patterns—that is, long-term distributional learning of the music of one's culture—may be necessary. But what is the role of learning in the perception of temporal patterns? Further, what musical elements (timbre, contour, rhythmic sequences) can be associated with metrical patterns such that those elements later affect processing of new musical instances? Creel (2011, 2012, 2013) has explored a phenomenon of top-down meter perception: listeners who hear melodies with ambiguous meter fill in meters based on their specific in-lab previous listening experiences. The major goal of the current paper is to specify what musical properties (timbre, contour, rhythmic sequences) influence this memory-based fill-in of meter. More broadly, such lab experiments offer an existence proof of influences of specific musical memories on perception of musical properties.

Previous research

A substantial body of literature on both music and linguistic prosody indicates that listeners use many cues other than the immediate auditory event to infer that event's meter and timing properties. These effects can broadly be classed into *recent context effects*—that is, influences of immediately-preceding context on meter perception; and *learning effects*, presumably due to more extensive, long-term exposure.

Recent context effects. Effects of context on meter(-like) processing are evident in both music and language. For instance, Povel and Essens (1985) found that listeners gave different assessments of rhythmic complexity to two identical rhythmic patterns when they were presented with different underlying beats, suggesting that the underlying beats formed part of the representation of the rhythm itself. Entrainment models (e.g. Large & Jones, 1999; Large, Herrera, & Velasco, 2015) further specify that listeners attune internal oscillators to metrical information, and the oscillator then continues at that period and phase, suggesting that recent preceding context is particularly important for processing timing (see also Barnes & Jones, 2000; Jones, Moynihan, Mackenzie, & Puente, 2002).

In the domain of language, Dilley and colleagues (e.g. Dilley & McAuley, 2008; Dilley, Mattys, & Vinke, 2010) have shown that listeners use *distal prosody* cues to segment words in running English speech. Specifically, they presented utterances of several syllables (channel dizzy foot-note-book-worm) and asked listeners to report the final word. The authors manipulated pitch and timing patterns of the first five syllables (channel dizzy foot…). Listeners reported different final words (*worm* vs. *bookworm*) depending on pitch and temporal properties of the first five syllables. This suggests some sort of stress-based entrainment of where the primary stress falls—not unlike beat entrainment seen in music. Thus, across domains, listeners seem sensitive to real-time contextual detail in processing auditory timing information. But do these real-time effects translate into long-term changes in processing?

Learning effects. Literatures on music and language processing both demonstrate learning effects on meter and timing perception. First, research on cross-cultural meter perception suggests a substantial learning (or unlearning) component (Hannon & Trehub, 2005a, 2005b; Hannon et al., 2012; Kalender et al., 2013). Infant listeners in Western cultures can detect

changes to complex, non-Western meters at 6 months, though Western adults cannot (Hannon &

Trehub, 2005a). At 12 months, Western infants can relearn these complex meters with exposure,

while Western adults do not (Hannon & Trehub, 2005b). Further studies suggest that complex-

meter-enculturated adults exceed Western adults in their perception of complex meters (Hannon

& Trehub, 2005a; Hannon et al., 2012; Kalender et al., 2013). One interpretation of this pattern

of results is that they reflect *learning* of meters in one's culture via massive exposure: the

acquisition of strong biases toward familiar meters. Whatever the interpretation, these results

suggest that metrical information may not simply be a property of the musical signal: it requires

either maintained sensitivity or learning.

Related phenomena occur in language. In particular, a number of studies suggest that

one's native language influences the ability to detect linguistic stress patterns: speakers of French

(which does not have metrical stress) have great difficulty in detecting stress patterns relative to

speakers of stress-timed languages such as German or English (e.g. Bhatara, Boll-Avetisyan,

Unger, Nazzi, & Höhle, 2013; Cutler, Mehler, Norris, & Segui, 1986). This may extend to timing

perception in music-like sequences (Bhatara, Boll-Avetisyan, Agus, Höhle, & Nazzi, 2016;

Iversen, Patel, & Ohgushi, 2008; Yoshida, Iversen, Patel, Mazuka, Nito, Gervain, & Werker,

2010).

Finally, some research suggests that meter can be learned as part of a melody with which

it cooccurs. Specifically, Creel (2011, 2012, 2013) presented listeners with melodies that were

ambiguous between 3/4 meter and 6/8 meter. During an exposure phase, a listener might hear

melody X with a 3/4 metrical accompaniment, and melody Y with a 6/8 metrical accompaniment.

The reverse might be true for another listener (X in 6/8, Y in 3/4). Then, in a test phase, all

listeners heard each melody without accompaniment. Their task was to judge how well a probe

series of drumbeats, presented immediately after the melody, fit with the melody itself. Listeners

provided higher goodness-of-fit ratings to probes that matched the meter they had heard during

the exposure phase. This suggests that they had associated metrical information, or at least

temporal properties of accompaniments, with the melodies themselves.

	Notation		Rhythmic pattern
V	♩	♫	4 1 1
W	♩.	♫♩	3 1 1 1
X	♩.	♫♩	3 1 1 1
Y	♩	♫♫	2 1 1 1 1
Z	♫♫♩		1 1 1 1 2

Figure 1. A sample of pitch contour + timing patterns, like those used by Creel (2012). While W
and X share rhythmic patterns, Creel (2012) treated them as different motifs because their
contours differed. Neither Creel (2012) nor the current study treat Y and Z as the same, even
though they are identical if phase (position of downbeat within the rhythm) is not considered.

Creel (2012) extended the Creel (2011) result in a series of experiments. Creel (2012)

presented listeners with multiple melodies, constructed so that all 3/4 melodies were heard in a

violin timbre, while all 6/8 melodies were heard in a clarinet timbre (or vice versa). After hearing

these melodies in the listening phase, participants heard the same melodies, unaccompanied, at

test. Melodies at test were heard either in their original timbre or in the opposite timbre. In these

cases, listeners gave higher goodness-of-fit ratings for the meter previously heard with that

melody than for the other meter, but only when melodies were heard in their original timbre

(Experiments 1 and 2). Learned meter associations did not generalize to new melodies with the

same timbre—that is, participants did not think a new clarinet melody necessarily fit better with 6/8 time. Further experiments asked whether brief rhythmic patterns with a particular contour (Figure 1), *motifs* for short, become associated with meter. To briefly summarize a complex set of findings, new melodies showed higher metrical fit ratings when they had both the same *timbre* and the same *motifs* as melodies heard at exposure. Thus, Creel's (2012) experiments suggest that *motif*-timbre-meter associations do generalize to new melodies, even though timbre-meter associations do not.

An exemplar account of music processing

Creel (2012) termed this metrical association or metrical fill-in phenomenon *metrical restoration*, on analogy with phoneme restoration (Warren, 1970; see also Samuel, 1981). This term is partly a notational convenience, as "metrical restoration" is a more compact descriptor than "preference for melody-specific familiarized meters."

However, it also reflects a particular perspective on musical listening and musical memory organization, specifically, an exemplar account of memory (Goldinger, 1998; Hintzman, 1986; Pierrehumbert, 2001). On an exemplar account, listeners store traces (exemplars) of each individual sound pattern they hear, similar to distributional learning accounts of auditory category formation (e.g. Maye, Werker, & Gerken, 2002). When a sound pattern is heard, it activates traces in memory *in proportion to their similarity to the input pattern* (Hintzman, 1986). That is, whenever someone is listening to music, that music activates previously-stored specific musical traces. Those activated traces join with the perceptual input to create a sense of tonality, time, and potentially other musical features.

In the current experimental context, if a particular melody is heard devoid of its previous accompaniment, the memory of the accompaniment may be reactivated. The sum of the reactivated memories, sometimes called an "echo" (Hintzman, 1986), then forms part of the percept itself. It is this memory echo that is postulated to generate metrical restoration. This similarity-based activation of specific instances contrasts with most models of music perception, which posit that the incoming trace activates abstract musical knowledge. For example, the Krumhansl-Schmuckler key-finding algorithm (see Krumhansl, 1990; Temperley, 1999), which compares incoming notes to general goodness-of-fit profiles of notes in a key, rather than comparing to particular musical pieces. Similarly, two recent beat-finding models (Large, Herrera, & Velasco, 2015; Tomic & Janata, 2008), wherein the musical input activates a set of oscillators rather than activating particular pieces of previously-heard music.

Exemplar models are appealing in two respects. First, they account for listeners' ability to store detailed musical information (e.g. Krumhansl, 2010; Schellenberg, Iverson, & McKinnon, 1999). Second, they provide a natural mechanism for generalization to novel instances—in the current case, novel melodies, via reactivation of specific, previously-heard musical instances. Thus they can account for specific knowledge and seemingly-abstract knowledge (via summation of multiple activated exemplars) in the same model.

While Creel's (2011, 2012) results are consistent with memory specificity, they provide somewhat less information about generalization to new instances. One way that generalization might occur is that, when one hears a new musical piece—say, a new jazz tune—the most strongly activated memories will be one or more similar jazz songs that one already knows. These activated memories sum with the perceptual input to generate a stylistic percept that includes meter information. If this *coactivation* of similar traces is really a major force in meter

processing, then melody-specific knowledge should at least sometimes generalize to other

melodies that are similar in style. This is to an extent a question about what counts as "similar":

the exemplar approach suggests that traces of *similar* music, not just identical traces, will be

coactivated and will shape meter processing. Thus, the goal of the current study is to explore

exactly what dimensions of similarity allow coactivation and thus affect generalization of

metrical processing.

It is noteworthy that the degree of metrical restoration was much greater in Creel (2011,

Experiment 2) than in any of the experiments in Creel (2012). While the differences in effect

magnitude between the two studies might reflect chance variation, they may hint that certain

musical features are more strongly conducive to meter induction than others. This question is

crucial for understanding how and when metrical memory patterns will generalize.

The current study

The present research constitutes an in-depth exploration of the influence of multiple

factors on metrical restoration (outlined in Table 1). Previous studies (Creel, 2011, 2012) suggest

that timbre, rhythmic patterns, and contour influence restoration. Yet multiple experiments

(Creel, 2012; unpublished data from my lab) suggest that a single cue, such as timbre, cannot

become associated with meter information on its own. Thus, the role of timbre uniqueness at the

individual melody level is explored here (Experiments 1 and 4).

Additionally, rhythmic patterns may play a role in metrical restoration. For example, in

Creel (2011)'s Experiment 2, which showed the strongest metrical restoration effects observed in

this paradigm to date, each melody was constructed primarily from its own single repeating

rhythmic pattern. On the other hand, in Creel (2012)'s Experiment 5, which showed only

moderate restoration effects, component rhythmic patterns were shared across melodies. This suggests that rhythmic patterns, apart from contour, may associate with metrical information. The role of rhythmic patterns is explored here (Experiments 2, 3, 4, 6), including amount of exposure to patterns, pattern similarity, pattern's (in)consistency of association with a particular meter, and immediate repetition of patterns, the last of which has been linked to meter perception (e.g. Temperley & Bartlette, 2002).

A third musical property that may become associated with meter information is *overall melodic form*—cues that are unique to a melody but do not include exact rhythmic patterns, such as the melody's coarse-scale contour. In the current study, to create melodies sharing overall melodic form, the original melody was modified by addition or subtraction of one note onset per measure, while maintaining overall melodic contour and implied harmonies (see Appendix C Figure C4 for examples of melodies sharing overall melodic form). Many researchers have inferred the existence of "pure" or "reduced" representations of a musical work, ranging from Schenker (1979) to Lerdahl and Jackendoff (1983). Such cues might include any of the reductions proposed by Lerdahl and Jackendoff (1983), including time-span reduction and prolongational reduction; large-scale contour (Morris, 1993); implied (or restored) harmony; or even *n*-grams of rhythmic patterns. Overall melodic form is one explanation for melody-specific metrical restoration effects in Creel's (2012) Experiments 1-3, where melodies overlapped substantially in rhythmic patterns. Effects of overall melodic form are investigated in Experiments 4, 5, and 6.

Table 1. Manipulations applied in various experiments.

Exp.	Features unique to melody	RP unique to	Swapped cues at test vs. meter

	Timbre	Rate	RP	RP + contour	meter	Match. RP	Confl. RP	Mat. timbre	Confl. Timbre
1	√.	√.	.	.	.	.	.	.	.
2	√	.	√	.	√	.	.	.	.
3	√	.	.	.	√.	.	.	.	.
4	√	.	√	.	√	√	√	√	√
5	√	.	√	.	√	.	√	.	√
6	√	.	.	√	√.	.	.	.	.

Note. RP = rhythmic pattern, Match. = matching, Confl., = conflicting, √ = element present, . = element absent, √. = element present for half of participants, absent for others.

Experiment 1: effects of timbre and rate diversity on metrical restoration

The first experiment explored the role of timbre diversity on listeners' ability to maintain particular meter associations with particular melodies. In exemplar terms, melodies that are farther apart in high-dimensional perceptual space are less likely to be activated by each other. This means that there should be less interference, such that only a particular melody's associated metrical information will be activated, leading to a clear metrical percept. The current experiment addressed this hypothesis by presenting 12 melodies either in unique timbres or in a single uniform timbre. Listeners were exposed to six 3/4 and six 6/8 melodies either in a single timbre, or in 12 different timbres (one per melody). If timbre diversity increases metrical restoration, then diverse-timbre learners should show stronger metrical restoration than single-timbre learners. To further increase melodies' perceptual uniqueness, the timbre-diversity manipulation was crossed with a rate-diversity manipulation: listeners heard melodies either at a single presentation rate, or in three different rates (consistent for a particular melody). If listeners store melodies rate-specifically, then multiple rates, like multiple timbres, should yield more distinct representations, and hence, stronger metrical restoration.

Method

Participants. Seventy-two listeners from the UCSD human participant pool took part in the experiment.

Stimuli. Stimuli were 12 major-mode melodies (examples in Appendix C) designed to include a range of rhythmic patterns within each melody (number of measures: $M = 8.5$, $SD = 1.0$; range: 8-11). The 12 timbres were selected to be discriminable (Iverson & Krumhansl, 1993) and to span multiple instrument families (percussion, strings, brass, woodwinds) representing a wide range of timbres. For a given participant, each melody had a single rate and timbre. Component rhythmic patterns were chosen to be moderately ambiguous with respect to 3/4 vs. 6/8 meter (see Appendix C). Rhythmic patterns occurred across multiple melodies.

Table 2. Experiment 1 example lists.

	Exposure trials											
	List 1			List 2			List 3			List 4		
Melody	Timbre	Rate	Meter	Timbre	Rate	Meter	Timbre	Rate	Meter	Timbre	Rate	Meter
1	violin	100	34	violin	100	68	vibraphone	120	34	vibraphone	120	68
2	muted tpt.	100	34	muted tpt.	100	68	pl. str.	120	68	pl. str.	120	34
3	clarinet	100	68	clarinet	100	34	flute	120	34	Flute	120	68
4	piano	100	68	piano	100	34	tpt.	120	68	tpt.	120	34
5	alto sax.	120	34	alto sax.	120	68	violin	140	34	violin	140	68
6	oboe	120	34	oboe	120	68	muted tpt.	140	68	muted tpt.	140	34
7	French horn	120	68	French horn	120	34	clarinet	140	34	clarinet	140	68
8	harp	120	68	harp	120	34	piano	140	68	piano	140	34
9	vibraphone	140	34	vibraphone	140	68	alto sax.	100	34	alto sax.	100	68
10	pl. str.	140	34	pl. str.	140	68	oboe	100	68	oboe	100	34
11	flute	140	68	flute	140	34	French horn	100	34	French horn	100	68
12	tpt.	140	68	tpt.	140	34	harp	100	68	harp	100	34
	Test trials											
Mel.	Timbre	Rate	Probe	Timbre	Rate	Probe	Timbre	Rate	Probe	Timbre	Rate	Probe
1	**violin**	**100**	**34**	violin	100	34	**vibraphone**	**120**	**34**	vibraphone	120	34
1	violin	100	68	**violin**	**100**	**68**	vibraphone	120	68	**vibraphone**	**120**	**68**
2	muted tpt.	100	34	**muted tpt.**	**100**	**34**	pl. str.	120	34	**pl. str.**	**120**	**34**
2	**muted tpt.**	**100**	**68**	muted tpt.	100	68	**pl. str.**	**120**	**68**	pl. str.	120	68

Note. **Bolded** test trials represent exposure-consistent responses. Tpt. = trumpet; pl. str. = plucked strings; sax. = saxophone.

Table 3. Experiment 1 melody and accompanying instrument pairings.

Melodic timbre	Accompanying timbre
violin	steel drums
muted trumpet	drawbar organ
clarinet	shamisen
piano	recorder
sax	brass section
oboe	harpsichord
French horn	guitar (nylon strings)
harp	accordion
vibe	piano
plucked string	church organ
flute	viola
trumpet (not muted)	music box

Design. Participants were randomly assigned to one of four combinations of timbre diversity (1 timbre or 12 timbres) and rate diversity (1 rate or 3 rates), 18 per group. Timbres and rates were counterbalanced such that each timbre, rate, and timbre-rate combination occurred roughly equally across participants and melodies. There were 12 unique lists per condition (examples of the full-variability condition in Table 2).

Procedure. The experiment was presented in Matlab using PsychToolBox 3 (Brainard, 1997; Pelli, 1997). Prior to the experiment proper, participants completed a set of questions on basic demographic information and prior music performing experience. Music experience was examined, but as it did not invalidate the reported results here, levels of music experience are described in Appendix A.

The experiment itself consisted of two distinct phases: an exposure phase, and a test phase. During the exposure phase, participants were simply asked to rate how much they liked the melody and how much they thought the melody sounded happy vs. sad. As in previous studies (Creel, 2011, 2012), this cover task was designed to discourage deliberate memorization of melodies' metrical properties. Listeners heard each exposure melody six times, embedded in a

metrical context (accompanying instruments that played along in a meter that clearly fit either

3/4 or 6/8; example in Figure 2). For a given melody and participant, the metrical context at

exposure was consistent—for instance, melody 1 for participant 2 was always heard in a 3/4

context. However, across melodies, each participant heard half 3/4 contexts and half 6/8

contexts; across participants, each melody was heard by half the participants in a 3/4 context and

by the rest in a 6/8 context.

Figure 2. Experiment 1, two example melodies (top lines) with a 3/4 context (lines 2-3) and a 6/8 context (lines 4-5). Only one context (that is, accompaniment) was heard at a time.
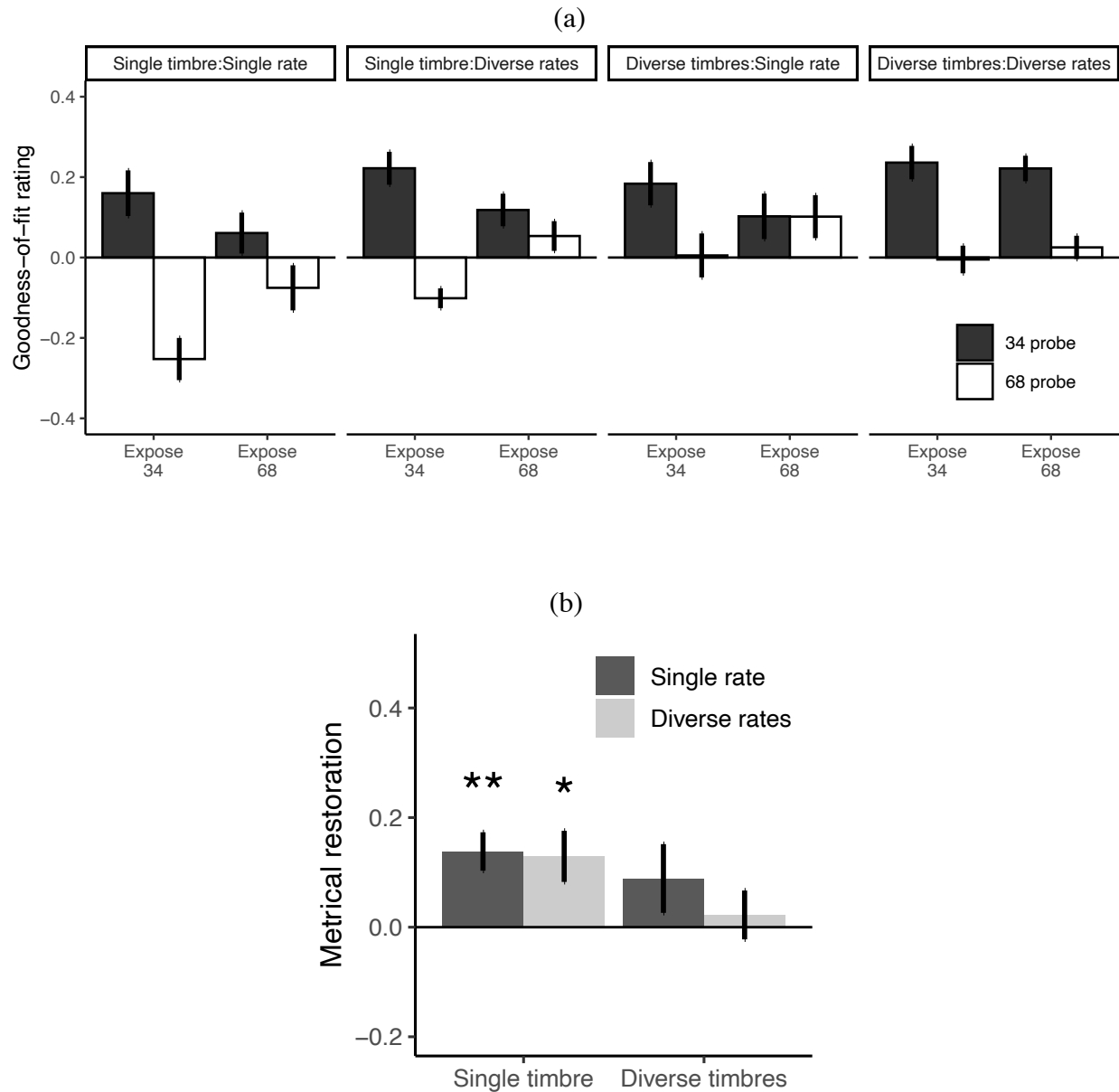
(a)

(b)

Figure 3. Experiment 1, metrical restoration effect as a function of timbre and rate variability. (a) Raw goodness-of fit ratings; (b) metrical restoration score. The first four bars in (a) correspond to the first bar in (b). Throughout, error bars are standard errors. ** *p* < .01, * *p* < .05

During the test phase, listeners were presented with each melody *devoid of accompaniment*, followed by a metrical "probe drumbeat" after the end of the melody in either 3/4 time or 6/8 time. They were asked to judge how well the drumbeat fit with the preceding melody. Four examples (two with "good" drumbeats, two with "bad" drumbeats) were provided

prior to the real test trials, using the tunes *Happy Birthday* (3/4) and *Greensleeves* (6/8). Then, across 48 total test trials, each listener heard each melody four times, twice with a 3/4 probe and twice with a 6/8 probe. Responses were made by clicking on a graphic of a ruler, which was 386 pixels in length and was labeled "Bad fit" at the left end, "Okay" in the center, and "Good fit" on the right end. The x-coordinate of the clicked pixel was transformed into a scale that varied from -1 (worst possible fit rating) to +1 (best possible fit rating).

Results

Metrical ratings are displayed in two ways. First, raw ratings (position on the ruler, where center = 0, worst possible fit = -1, best possible fit = 1) are displayed in Figure 3a and Appendix D. I take as evidence for metrical restoration higher ratings for the previously-exposed meter for that melody than for the non-exposed meter. Statistically, this shows up as an Exposure Meter (3/4, 6/8) x Probe Meter (3/4, 6/8) interaction. In practice, the higher ratings for exposure-consistent metrical probes may be superimposed on an overall preference for a particular meter, usually 3/4 (see Creel, 2012), perhaps because of listeners' general preference for beats with two subdivisions rather than three (e.g. Parncutt, 1994). This pattern (metrical restoration plus a 3/4 preference) is evident in the first two conditions depicted in Figure 3a. The fourth (rightmost) condition in Figure 3a shows only an overall preference for 3/4 probes, without detectable evidence of metrical restoration. Since the depiction in Figure 3a is quite dense, the overall restoration effect is summarized in a single value for each condition in Figure 3b. Mathematically, this is simply the average fit ratings for exposure-matching probes minus the average fit ratings for mismatching probes (see Appendix D). Notice that this formulation

cancels out *overall* preferences for one meter or the other. For simplicity, further data figures follow the pattern of Figure 3b, with full data patterns in Appendix D.

Metrical restoration was assessed in an ANOVA on probe ratings, with Exposure Meter and Probe Meter as within-participants and within-items factors, and Timbre Diversity (one or twelve) and Rate Diversity (one or three) as between-participants (within-items) factors. For simplicity, throughout, only effects involving the Exposure Meter x Probe Meter are reported, because this interaction indicates metrical restoration. Effect sizes are reported as generalized eta-squared (Bakeman, 2005; Olejnik & Algina, 2003).

An Exposure Meter x Probe Meter interaction ($F1(1,68) = 14.52, p = .0003$; $F2(1,11) = 4.81, p = .05$; $\eta_G^2 = .02$) indicated a metrical restoration effect overall. However, none of the higher-level interactions—which would indicate timbre diversity or rate diversity effects—were significant: Exposure Meter x Probe Meter x Timbre Diversity ($F1(1,68) = 2.47, p = .12$; $F2(1,11) = 1.74, p = .21$; $\eta_G^2 = .00$), Exposure Meter x Probe Meter x Rate Diversity ($F1(1,68) = 0.58, p = .45$; $F2(1,11) = 0.35, p = .57$; $\eta_G^2 = .00$), or Exposure Meter x Probe Meter x Timbre Diversity x Rate Diversity ($F1(1,68) = 0.34, p = .56$; $F2(1,11) = 0.60, p = .45$; $\eta_G^2 = .00$). Further, as is visible in Figure 3b, the direction of the Timbre Diversity effect was numerically opposite that predicted: metrical restoration was more robust for listeners who heard a single timbre than for those who heard 12 different timbres. Individually, metrical restoration was significant only in the two single-timbre conditions (no-variability: $F1(1,17) = 14.64, p = .001$; $F2(1,11) = 4.47, p = .06$; $\eta_G^2 = .03$; rate-variability: $F1(1,17) = 7.25, p = .02$; $F2(1,11) = 7.07, p = .02$; $\eta_G^2 = .07$). Thus, the strongest evidence for metrical restoration was carried by the low-timbre-diversity conditions.

Discussion

The results of Experiment 1 were counter to predictions of timbre specificity and rate specificity. Instead of stronger metrical restoration when each melody had a unique timbre, metrical restoration was numerically smaller—and absent—when each melody had a unique timbre. Further, there was no effect of rate diversity, even though one might think that rate would be closely linked to meter as both are aspects of musical timing. This effect is somewhat surprising: why wouldn't 12 more-distinct melodies (12 unique timbres) generate stronger metrical restoration, due to greater perceptual uniqueness of the melodies, than 12 less-distinct (identical-timbre) melodies?

One possible answer is that listeners were not associating wholesale timbres with meter, but were associating *timbre-specific rhythmic patterns* with meters (e.g., 2 1 1 1 1 in clarinet timbre). That is, listeners were aggregating exemplars that grouped according to perceptual similarity. Recall that Creel (2012) showed that sets of particular rhythm-plus-contour elements influenced metrical restoration. Suppose that listeners in the current experiment also associated meter with such patterns. The melodies in the current study were built from a small set of moderately-ambiguous rhythmic patterns, many of which occurred across multiple melodies. If the same rhythmic patterns were stored separately for separate timbres, then listeners who heard a single timbre might build up relatively stronger pattern representations. On the other hand, listeners who heard 12 timbres would store a larger number of timbre-specific exemplars, but each of these would be weaker because they had been experienced far fewer times, leaving them unable to generate significant metrical restoration. Of course, all of this supposition is grounded in a null effect of timbre diversity, so caution is warranted.

It should also be acknowledged that effects may be driven by similarity relations between the timbres and rates tested. That is, perceptually-similar timbres or rates may have interfered with each other, diluting any effects of timbre or tempo specificity. The study also used a rather restricted range of rates. Perhaps if more-distinct tempos had been used, rate effects would have been evident, though outside of a limited range of rates, listeners might perceive timing patterns at a different level of the metrical hierarchy (see, e.g., London, 2002; Parncutt, 1994).

These data may also imply that rhythmic patterns are particularly critical to restoration: when a particular rhythmic pattern is scattered across melodies and meters, restoration effects are overall relatively weak. Thus, the next experiment turned to considering the role of rhythmic pattern uniqueness in metrical restoration. Like the diverse-timbres condition of Experiment 1, each melody had a unique timbre. However, unlike Experiment 1 (but like Creel, 2011, Experiment 2, which showed strong restoration effects), each melody also had its own unique, repeating rhythmic pattern.

Experiment 2A&B: amount of exposure to, and similarity between, rhythmic patterns

This experiment addressed the hypothesis that the strongest metrical restoration effects come as a result of consistent mappings between a particular meter and a particular rhythmic pattern. That is, each rhythmic pattern occurs in only *one* of the two possible metrical settings. The current experiment presented 8 melodies, each of which had its own timbre and its own unique repeating rhythmic pattern. Thus, it was like Experiment 1 in using diverse timbres, but differed from Experiment 1 in using distinct rhythms. If unique rhythms are important, then metrical restoration should increase in magnitude relative to Experiment 1.

Table 4. Experiment 2, rhythmic patterns and timbres.

Melody	Rhythmic pattern	Timbre (melody)	Timbre (context)
1	4 1 1	Piccolo	String ensemble
2	3 1 1 1	Clarinet	Flute, string ensemble
3	1.5 .5 1 1 2	Vibraphone	Recorder, bassoon
6	1 1 1 1 2	Guitar (nylon strings)	French horn
4	2 1 1 2	English horn	Pizzicato strings, contrabass
5	3 1 2	Harpsichord	Cello
7	2 1 1 1 1	Trumpet	Trombone
8	1 1 1 1 1 1	Viola	Church organ

Note. Pairs of melodies denoted by shading were classed as similar rhythmic patterns.

Two additional factors were manipulated. First was number of exposures. Listeners heard each melody between two times and sixteen times during the exposure phase. If metrical restoration effects require high quantities of exposure, then effects should show up more strongly with more repetitions. The second additional factor manipulated was match or mismatch in accompanying meter between melodies with similar rhythmic patterns. Similar rhythmic patterns are depicted sequentially in Table 4. For example, a listener in the "patterns match" condition might hear the 4 1 1 melody in 3/4 time, and the 3 1 1 1 melody in 3/4 time, while a listener in the "patterns conflict" condition might hear the 4 1 1 in 3/4 but the 3 1 1 1 in 6/8. "Similar" here was defined as pairs with an edit distance of 1—that is, one addition, deletion, or temporal change of a single onset time: the 4 1 1 pattern can be changed to the 3 1 1 1 pattern by adding an onset on the fourth sub-beat. (Note that there are other possible similarity relationships amongst this set as well, but on average, each pattern only matched about 1 additional pattern besides its paired pattern.) If traces of similar, but not identical, rhythmic patterns are coactivated during listening, then listeners who hear conflicting meters for similar rhythmic patterns might show smaller metrical restoration effects than listeners who hear matched meters for similar rhythmic patterns.

Method

Participants. N=48 participants took part in Experiment 2A. Half heard each melody 4 times during exposure, the other half heard each melody 8 times during exposure. A later group of 48 participants (Experiment 2B) heard each melody either 2 times (24 participants) or 16 times (24 participants) during exposure, to explore differences in amount of exposure.

Stimuli. A set of eight new melodies was created (examples in Appendix C). Each melody was dominated by a single measure-long rhythmic pattern and a unique timbre. Each melody also had an A-A' structure: the first (A) phrase was followed by a similar (A') phrase. Each melody was roughly 16 measures long ($M = 16$, $SD = 0.5$). The rhythmic-pattern design of the melodies meant that each individual rhythmic pattern occurred almost exclusively with a single meter for a given participant—either 3/4 or 6/8, but not both.

Procedure. The procedure matched that of the previous experiment, except for the changes to stimuli, lists, and number of presentations per melody.

Design. Participants heard 2, 4, 8, or 16 repetitions of each melody during exposure. For 16-repetition participants only—who heard 128 exposure trials—a one-minute break was provided halfway through exposure due to the extensive nature and repetitiveness of the stimuli.

For each participant, melodies were presented in random order without replacement in blocks of 8 melodies. Each listener heard four of the melodies with contexts in 3/4 time, and the other four in contexts in 6/8 time. This was counterbalanced across listeners (Table 5) so that each melody was equally likely to occur in each time signature. At test, each melody was presented without a context four times: twice followed by a 3/4 drumbeat probe, and twice followed by a 6/8 drumbeat probe, yielding 32 test trials total.

Table 5. Experiment 2, different assignments of melodies to meters.

Melody	List											
	1	2	3	4	5	6	7	8	9	10	11	12
1	34	68	34	68	34	68	34	68	34	68	34	68
2	34	68	34	68	34	68	68	34	68	34	68	34
3	34	68	68	34	68	34	34	68	34	68	68	34
6	34	68	68	34	68	34	68	34	68	34	34	68
4	68	34	34	68	68	34	34	68	68	34	68	34
5	68	34	34	68	68	34	68	34	34	68	34	68
7	68	34	68	34	34	68	34	68	68	34	34	68
8	68	34	68	34	34	68	68	34	34	68	68	34

Note. Lists 7-12 presented similar-rhythmic-pattern melodies with conflicting meters. For instance, melodies 1 and 2, which had similar rhythmic patterns, were presented during exposure with different meters.

Results

Listeners overall showed robust metrical restoration effects (Figure 4). Mixed ANOVAs with Exposure Meter and Probe Meter as within-participants factors, and Amount of Exposure and Pattern Conflict as between-participants (within-items) factors, were run on probe ratings. There was a strong Exposure Meter x Probe Meter interaction ($F1(1,88) = 43.98, p < .0001$; $F2(1,7) = 15.95, p = .005$; $\eta_G^2 = .13$), indicating a metrical restoration effect overall. There was a marginal interaction of Pattern Conflict x Exposure Meter x Probe Meter ($F1(1,88) = 2.98, p = .09$; $F2(1,7) = 2.39, p = .17$; $\eta_G^2 = .01$). This implies that there may be a slight amount of interference between similar rhythmic patterns, but if so, it is not a strong effect. Finally, the interaction of Amount of Exposure x Exposure Meter x Probe Meter did not approach significance ($F1(1,88) = 1.73, p = .17$; $F2(3,21) = 4.00, p = .02$; $\eta_G^2 = .02$), suggesting that the visible differences in Figure 4 were not reliable. No other effects approached significance.

Despite the null effect of Amount of Exposure, the effects at each level of exposure were examined to test whether they were reliable individually. For 2-exposure participants, Exposure Meter x Probe Meter was significant ($F1(1,23) = 4.79, p = .04$; $F2(1,7) = 6.94, p = .03$;

$\eta_G^2 = .07$), suggesting that as few as two exposures to a melody suffices to generate metrical

restoration (note that there were on average 11 repetitions of the rhythmic pattern in each melody

[range: 8-14]). The 4-exposure participants ($F1(1,23) = 8.35, p = .008; F2(1,7) = 4.87, p = .06;$

$\eta_G^2 = .07$), 8-exposure participants (F1(1,23) = 22.24, p < .0001; F2(1,7) = 19.35, p = .003;

$\eta_G^2 = .26$), and 16-exposure participants ($F1(1,23) = 12.71, p = .002; F2(1,7) = 14.13, p = .007;$

$\eta_G^2 = .10$) all showed significant restoration effects as well. Further, a comparison of the overall

metrical restoration effects in Experiments 1 and 2 indicated higher-magnitude restoration in the

current experiment ($t(161.6) = 3.10, p = .002$, Cohen's $d = .49$). This is consistent with more

robust metrical restoration when rhythmic patterns are consistently mapped to a melody

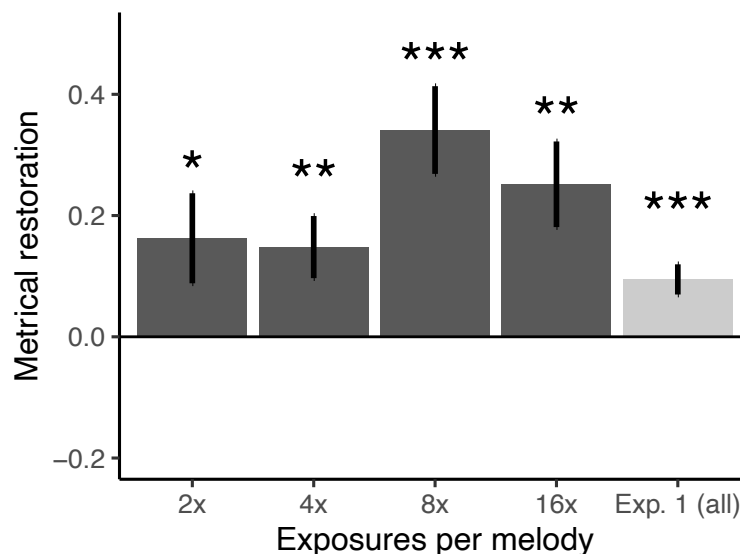(Experiment 2) than when they are not (Experiment 1).

Figure 4. Experiment 2, metrical restoration (dark gray), with overall effects from Experiment 1
for comparison (light gray). * $p < .05$, ** $p < .01$, *** $p < .001$

Of course, a counterexplanation of the greater metrical restoration in Experiment 2 is that

something about the melodies in Experiment 2 was more malleable to meter manipulation than

those in Experiment 1. In particular, perhaps Experiment 1 melodies in isolation were less

metrically ambiguous, leaving less room for accompanying meters to have an effect. This was

tested in a control experiment, described fully in Appendix B. In brief, no metrical properties of

the Experiment 1 vs. Experiment 2 melodies differed in ways that could plausibly generate

stronger restoration effects in Experiment 2.

Discussion

There was stronger metrical restoration in the current experiment than in Experiment 1.

While melodies in both experiments were timbre-diverse, only the melodies in the current

experiment contained consistent mappings between meters and rhythmic patterns, and only the

current experiment showed strong restoration effects. This suggests that timbre diversity does not

decrease metrical restoration, but that weak associations between rhythmic patterns and meter do.

Timbre *consistency* effects—stronger restoration when melodies share timbres, as numerically

evident in Experiment 1 (see also Creel, 2012 Experiments 1-2)—may surface more strongly

when rhythm-meter associations are weak. That is, when rhythm-meter mappings are

inconsistent, the influence of timbre consistency may be relatively greater.

Impressively, listeners showed significant metrical restoration after only two exposures to

a melody (on average, 22 instances of hearing a particular rhythmic pattern in a particular meter).

However, a curious question is why there was not a significant effect of exposure. Examining

Figure 4 suggests that the 2-exposure and 4-exposure conditions showed numerically weaker

restoration than the 8-exposure and 16-exposure conditions, though the 16-exposure condition

showed slightly weaker restoration relative to 8 exposures. This might indicate that the 8-

exposure condition was anomalously high due to chance variation. It might instead indicate that

the 16-exposure participants were experiencing fatigue from the unusually long duration of the experiment.

Another question was whether there was competition or interference between similar rhythmic patterns. That is, if two melodies with similar rhythmic patterns had different meters, would these coactivate and interfere with each other, weakening metrical restoration? The answer seems to be no: Across all participants, this was a marginal effect, suggesting that interference between similar but nonidentical rhythmic patterns is relatively minimal.

The current experiment suggests that rhythmic patterns play a strong role in metrical restoration. However, multiple accounts can explain these data. First, listeners may be associating particular rhythmic patterns with particular meters. Second, they may be associating a particular rhythmic pattern with a particular *melody*, perhaps requiring that a single rhythmic pattern cycle repeatedly within a melody (Temperley & Bartlette, 2002). Third, timbre cues may be facilitating metrical restoration here (even though they did not in Experiment 1) because timbres here are uniquely associated with a rhythmic pattern.

Leaving aside the timbre question for the time being, the next experiment addresses the first two accounts: associating rhythmic patterns with meters vs. associating rhythmic patterns with a particular melody. To test this, similar rhythmic patterns were shared across pairs of melodies. That is, each melody consisted of alternations between two similar rhythmic patterns (4 1 1, 3 1 1 1, …), and another melody used those same two patterns (3 1 1 1, 4 1 1, …). Half of the participants heard these pairs of melodies in the same meter (patterns match condition), the other half heard them in different meters (patterns conflict condition). Linking this to the two accounts stated above, if listeners associate rhythmic patterns with meters, then there should be strong restoration in the patterns-match condition, but weak or no restoration in the patterns-

conflict condition. If instead listeners associate particular melodies with rhythmic patterns, or if they require immediate repetitions of a brief rhythmic pattern with a melody (Temperley & Bartlette, 2002), then metrical restoration should be weak in both conditions because rhythmic patterns are not consistently associated with a melody.

Experiment 3: melodies share rhythmic patterns

Here, pairs of melodies from Experiment 2 were modified so that each melody in the pair contained two similar rhythmic patterns (examples in Appendix C). Varied across participants, the two melodies in a pair received either *conflicting* meters during exposure, or *matching* meters during exposure. If unique associations between a rhythmic pattern and a meter are important to metrical restoration, then only participants in the *matching* condition will show metrical restoration. If association with a unique melody is sufficient, then both *matching* and *conflicting* conditions should show metrical restoration.

Method

Participants. *N*=48 new participants took part.

Stimuli. These were the same as in Experiment 2, except for the rearrangement of motifs.

Procedure. This matched Experiment 2, with the exception that only 4-exposure and 8-exposure conditions were run.

Results

Figure 5 suggests a strong effect of rhythmic pattern conflict, with no restoration for the conflicting rhythmic patterns condition, but very strong restoration in the matching rhythmic

patterns condition. To assess whether rhythmic pattern conflict and number of exposures affected

metrical restoration, ANOVAs were conducted on participants and items with Exposure Meter

(34, 68), Probe Meter (34, 68), Amount of Exposure (4 per melody, 8 per melody) and Pattern

Conflict (conflict, match) as factors. Amount of Exposure and Pattern Conflict were between-

participants factors; all others were within participants and items.
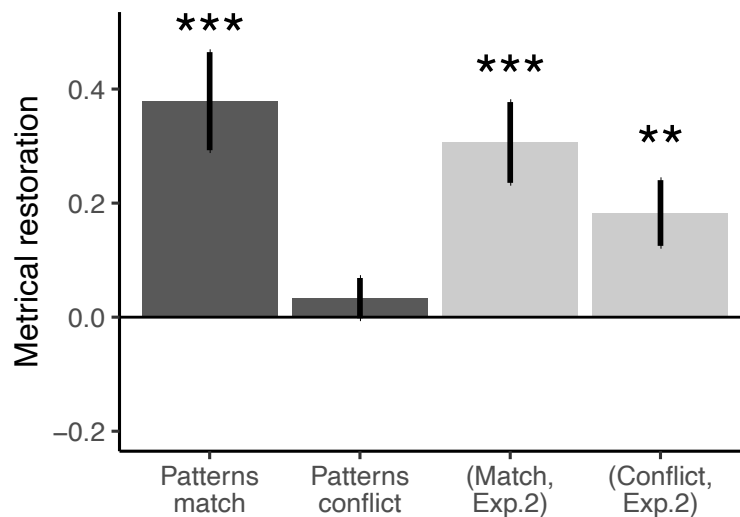

Figure 5. Experiment 3 metrical restoration (dark gray), with Experiment 2's analogous
conditions (light gray). *** $p < .001$, ** $p < .01$

There was an effect of Exposure Meter x Probe Meter ($F1(1,44) = 19.79, p < .0001$;

$F2(1,7) = 24.57, p = .002$; $\eta_G^2 = .09$). This resulted from an overall tendency to rate metrical

probes higher when they matched the exposure meter. This interaction was qualified by a higher-

level interaction of Pattern Conflict x Exposure Meter x Probe Meter ($F1(1,44) = 13.90, p$

$= .0005$; $F2(1,7) = 10.54, p = .01$; $\eta_G^2 = .06$), suggesting different patterns of Exposure Meter x

Probe Meter at each level of Pattern Conflict. Simple ANOVAs considering each level of Pattern

Conflict revealed a robust Exposure Meter x Probe Meter interaction for pattern-matched

melodies ($F1(1,23) = 19.40, p = .0002$; $F2(1,7) = 31.30, p = .0008$; $\eta_G^2 = .26$), while there was

no interaction for pattern-conflict melodies ($F1 < 1$; $F2 < 1$). That is, metrical conflict between

melody pairs containing the same rhythmic patterns erased the metrical restoration effect. No

other effects approached significance, including Exposure Meter x Probe Meter x Amount of

Exposure, suggesting weak effects of amount of exposure.

To assess whether immediate rhythmic pattern repetition affected restoration, the current

study's pattern match condition (where rhythmic patterns alternated within a single melody) was

compared with Experiment 2's pattern match condition (a single pattern within a melody; 4 and

8 exposure conditions only, to match with the exposure conditions used in the current

experiment). This ANOVA amounted to comparing the first and third bars in Figure 5. There

was a strong effect of Exposure Meter x Probe Meter ($F1(1,44) = 38.61$, $p < .0001$; $F2(1,7) =$

21.31, $p = .002$, $\eta_G^2 = .26$), consistent with the strong effect of metrical restoration overall.

However, the interaction of these factors with Experiment was far from significant ($F1(1,44) =$

0.43, $p = .52$; $F2(1,7) = 0.33$, $p = .59$, $\eta_G^2 = .00$), suggesting that the absence of immediate

rhythmic pattern repetitions in Experiment 3 did not degrade metrical restoration relative to

Experiment 2.

Discussion

This experiment explored the role of rhythmic consistency within a melody. There

appears to be no harm in sharing rhythmic patterns with another melody, as long as that melody

has the same meter: when two melodies shared two rhythmic patterns, restoration was as strong

as in Experiment 2, where each melody had its own rhythm. However, complete rhythmic

pattern overlap between two melodies with opposing meters canceled out metrical restoration

entirely. Relatedly, this experiment assessed whether restoration is stronger when rhythmic

patterns occur cyclically (immediately repeating themselves). There appears to be no restoration benefit for repeating a single measure-long rhythmic pattern (Experiment 2) over a pair of patterns (Experiment 3). Complementing Temperley and Bartlette's (2002) findings, this may indicate that while pattern repetitions are relevant for detecting events at the period of the repeating pattern itself (here, full measures, not beats), they do not carry information about the metrical subdivisions of that period (that is, whether the full measure is in $2+2+2 = 3/4$ or in $3+3 = 6/8$).

These results are somewhat consistent with Creel's (2012) finding that *rhythm plus contour* patterns (referred to in that study as motifs) generate restoration. However, it differs in that the current study explores the role of rhythmic patterns alone, with the particular contour of the pattern varying across instances.

The current experiment emphasizes the strength of rhythmic pattern-meter associations. However, rhythmic patterns did not change from exposure to test—test melodies all contained their original rhythmic patterns. Thus, it does not tell us much about how rhythmic patterns shape *generalization* of metrical information to *new* music, nor does it dissociate the role of specific rhythmic patterns from overall melodic form. The remaining experiments explore the role of overall melodic form vs. rhythmic patterns in influencing generalization to new musical materials. The guiding question: do brief rhythmic patterns almost completely dictate metrical restoration effects, or do other, coarse-grained melodic form cues play a role as well? Experiment 4 examines whether changes in rhythmic pattern and timbre from exposure to test affect metrical restoration. Experiment 5 focuses on conflict between rhythmic patterns and melodic form. Finally, Experiment 6 assesses whether pitch contour, in combination with rhythmic pattern, has strong effects on metrical restoration.

Experiment 4: changing cues to meter after learning

The current experiment began to explore whether rhythmic patterns solely shape metrical restoration, or instead whether other cues—timbre, melodic form—also work to shape restoration. More specifically, if Melody A is learned in 3/4 meter and flute timbre, but is then heard at test with a rhythmic pattern and/or timbre linked to 6/8 meter, will listeners prefer 3/4 time (matching the melodic form), or 6/8 time (matching the rhythmic pattern and/or timbre)?

The experiment was carried out with a modified version of the rhythmic pattern match/conflict materials from Experiment 2 (examples in Appendix C). For all listeners, half of test melodies were heard with their original rhythmic patterns, and half were heard with rhythmic patterns swapped in from the paired similar-pattern melody. Importantly, if simply changing the rhythmic patterns in a melody from exposure to test interferes with metrical restoration, then all listeners should show weaker metrical restoration in swapped-rhythmic pattern than in original-rhythmic pattern melodies. However, if changing the rhythmic patterns only interferes when the new rhythmic pattern has the opposite meter of the original melody, then only the conflicting-rhythmic pattern participants should show weakened—or perhaps even reversed—metrical restoration for swapped-rhythmic pattern melodies.

Method

Participants. $N = 48$ participants took part.

Stimuli. The melodies from Experiment 2 were modified by crossing the timbre and rhythmic patterns of similar-rhythmic pattern melody pairs. Thus, melodies 1 and 2 yielded 8

combinations (Table 6). Melodies with different rhythmic patterns are shown in Appendix C.

Different participants were trained on each of the four melody pairs shown in Table 6.

Table 6. Experiment 4, example of timbre and rhythmic pattern rearrangement, melodies 1-2, with example participant exposure assignments.

Participant	Melody	Timbre	Rhythmic Pattern
A	1	Piccolo	4 1 1
	2	Clarinet	3 1 1 1
B	1	Clarinet	4 1 1
	2	Piccolo	3 1 1 1
C	1	Piccolo	3 1 1 1
	2	Clarinet	4 1 1
D	1	Clarinet	3 1 1 1
	2	Piccolo	4 1 1

Design. Half of participants were exposed to melodies such that similar-pattern melodies (e.g. melody 1 and melody 2) had the same meter. The other half of participants were exposed such that similar-pattern melodies had opposite meters. Each listener heard 8 melodies 8 times each during exposure (total of 64 exposure trials). All participants were tested on the same set of melodies. This meant that for half of the participants, all melodies were tested in the original timbre, and for the other half, all melodies were tested in the opposite timbre. For each participant, half of the tested melodies were original-pattern and half were swapped-pattern melodies.

Results

Figure 6 suggests an interaction between pattern conflict and pattern swapping. ANOVAs with Pattern Conflict (match, conflict), Timbre Swapping (original, swapped; between-participants), Pattern Swapping (original, swapped), and Exposure Meter x Probe Meter were

conducted. The Exposure Meter x Probe Meter interaction was significant ($F1(1,44) = 11.49, p$

$= .001; F2(1,7) = 9.90, p = .02; \eta_G^2 = .05$), indicating the presence of metrical restoration overall.

This interaction was qualified by a three-way Exposure Meter x Probe Meter x Pattern Conflict

interaction ($F1(1,44) = 4.93, p = .03; F2(1,7) = 20.57, p = .003; \eta_G^2 = .02$), which resulted from

an overall larger metrical restoration effect when the exchanged rhythmic patterns matched in

meter than when they conflicted in meter. The Exposure Meter x Probe Meter x Pattern Conflict

x Pattern Swapping interaction was marginal by participants ($F1(1,44) = 3.17, p = .08; F2(1,7) =$

$2.35, p = .17; \eta_G^2 = .01$), indicating a tendency for different metrical restoration effects as a

function of whether similar patterns—the ones that switched between melodies—had been

learned with the same meter or two different meters. No other factors involving the Exposure

Meter x Probe Meter interaction approached significance, including those involving Timbre
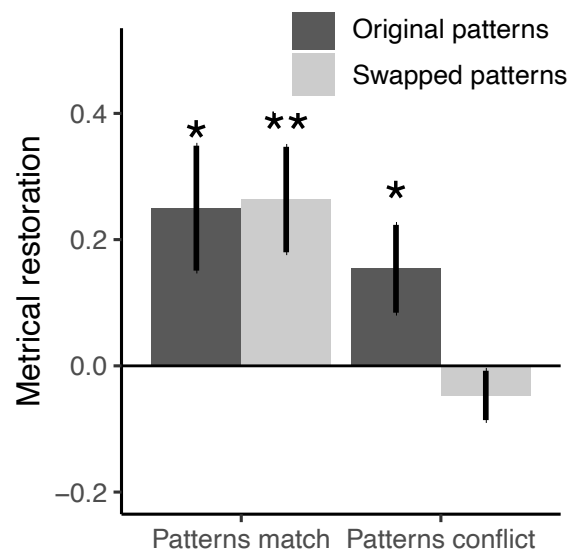
Swapping.

Figure 6. Experiment 4, metrical restoration. ** $p < .01$, * $p < .05$

Because there was a specific prediction that the effect of pattern swapping would disproportionately affect the metrically-conflicting-patterns condition, the two levels of Pattern Conflict were analyzed separately, collapsing over Timbre Swapping. For the matching-patterns condition, only Exposure Meter x Probe Meter was significant ($F1(1,23) = 10.13$, $p = .004$; $F2(1,7) = 23.66$, $p = .002$; $\eta_G^2 = .11$). This implied that metrical restoration was equally strong when a melody was heard with its original pattern, or with a pattern from another melody originally heard in the same meter: just changing the rhythmic pattern in a melody does not necessarily impede metrical restoration.

However, for the conflicting-patterns condition, there was a three-way Pattern Swapping x Exposure Meter x Probe Meter interaction (significant by participants; $F1(1,23) = 6.33$, $p = .02$; $F2(1,7) = 3.43$, $p = .11$; $\eta_G^2 = .02$). Breaking down the data by Pattern Swapping (original vs. swapped) showed that the Exposure Meter x Probe Meter interaction was significant only for the original-pattern melodies (significant by participants; $F1(1,23) = 4.88$, $p = .04$; $F2(1,7) = 5.02$, $p = .06$; $\eta_G^2 = .05$), but not for the swapped-pattern melodies ($F1(1,23) = 1.44$, $p = .24$; $F2(1,7) = 0.23$, $p = .65$; $\eta_G^2 = .01$). This implies that changing a melody's rhythmic patterns to ones originally heard in the other meter (implanting 34 patterns into a 68 meter, or vice versa) obliterates the metrical restoration effect. Interestingly, though, it does not significantly *reverse* the metrical restoration effect, suggesting that rhythmic patterns may not drive the effect in isolation.

Discussion

The current experiment asked whether rhythmic patterns alone drove metrical restoration effects, or if other factors impinge. Changing the rhythmic patterns of a melody only weakened

restoration when the replacement rhythmic pattern was associated with the opposite meter. Thus, just changing the component rhythms in a melody did not necessarily decrease restoration effects (for instance, by making the melody less recognizable)—the replacement pattern had to be associated with a different meter to weaken metrical restoration. This suggests that some metrical properties can generalize across a change in the rhythms of a melody. It also suggests either that rhythmic pattern effects conflict with overall melodic form effects, or that the change in rhythmic pattern affects melody recognition much more for the pattern-conflict condition than for the pattern-match condition. The between-participants timbre manipulation did not yield significant changes in restoration, suggesting that—within the current set of stimuli—rhythmic patterns and overall melodic form have more impact on restoration, while restoration generalizes over timbre.[1]

The question remains as to how changing the rhythmic pattern in a melody to a pattern that mismatches the melody-associated meter weakens metrical restoration. Is it because changing the rhythmic pattern makes the melody less recognizable, leading to only weak activation of the original melody's meter? Instead, is the pattern-swapping effect due to active interference between pattern-meter and melody-meter properties, as an exemplar account would predict? If the rhythmic pattern is controlling the metrical percept, then the metrical restoration effect should *reverse* when the rhythmic pattern changes to one associated with the opposite meter. It is difficult to say from Experiment 4 which of these is the case. The metrical restoration effect is numerically, but not significantly, in the wrong direction in the pattern-conflict

---

[1] Since previous studies (Creel, 2012, Experiments 1-2; Creel, 2013) have shown timbre effects on metrical restoration, a control experiment was run to replicate the timbre switch from the current experiment. Listeners heard exposure trials as in Experiment 4, but in the test phase, heard rhythmically-unaltered melodies that either had the original timbre or the changed timbre. Metrical restoration was equivalent regardless of timbre change, suggesting that, for the current materials, timbre does not strongly influence metrical restoration.

swapped-pattern melodies, but this was only tested on half of the participants, providing less

statistical power. Therefore, the next experiment examined the interplay between rhythmic

pattern and melodic form more closely with all participants receiving melodic form/rhythmic

pattern conflict trials.

Experiment 5: changing rhythmic patterns after learning

This experiment assesses whether switching the rhythmic pattern of a familiarized

melody to one associated with the other meter completely reverses the meter percept. That is, if

Melody 1 was heard in 3/4 during exposure, but is played at test with rhythms from a different,

6/8 melody, will the percept of Melody 1 now be completely like 6/8? Or does the overall form

of Melody 1 still exert some effects? To assess this, all listeners received exposure as in the

pattern-conflict conditions of the Experiment 4. They were then all tested on both original-

patterns-original-timbre melodies, and swapped-patterns-swapped-timbre melodies. The full

originals were included during test to maintain active traces of the original pattern-meter and

melody-meter mappings.

Method

Participants. $N = 48$ participants took part.

Stimuli and Procedure. The conflicting-pattern conditions of Experiment 4 were used as

exposure stimuli. Each participant heard 8 melodies 8 times each during exposure, for a total of

64 exposure trials. The test, however, contained two types of stimuli: either the *full original*

timbre and rhythmic pattern as in exposure, or a *full swap* in timbre and pattern. The logic was

that a larger number of participants combined with maintained full match trials during test would

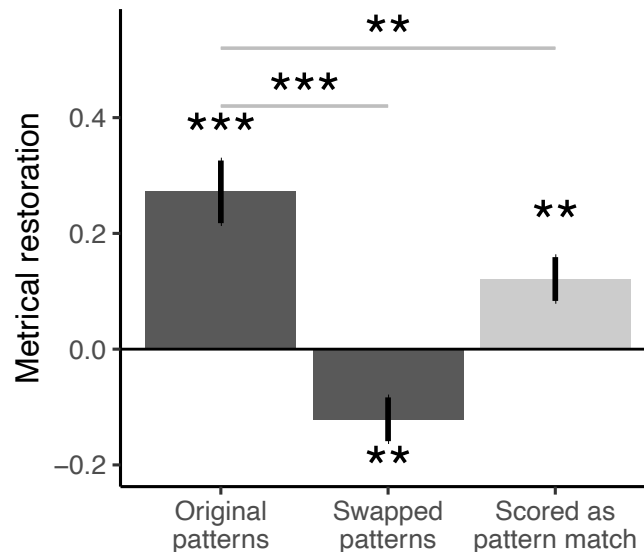reveal whether a change in rhythmic pattern could fully reverse the metrical percept.

Figure 7. Experiment 5, metrical restoration effect. *** $p < .001$, ** $p < .01$

Results

Figure 7 suggests a strong effect of swapping rhythmic patterns on metrical restoration.

ANOVAs were conducted with Pattern Swapping (original, swapped), Exposure Meter, and

Probe Meter as within-participants and within-items factors. The Exposure Meter x Probe Meter

interaction was significant ($F(1,47) = 8.06, p = .007; F(1,31) = 6.93, p = .01; \eta_G^2 = .01$),

indicating metrical restoration overall. The three-way interaction was also significant ($F(1,47) =$

$26.18, p < .0001; F(1,31) = 41.87, p < .0001; \eta_G^2 = .07$), indicating different Exposure Meter x

Probe Meter effects depending on whether listeners heard a melody with its original pattern and

timbre or with the swapped pattern and timbre. To examine the three-way interaction, simple

ANOVAs with Exposure Meter and Probe Meter as factors were conducted at each level of

Pattern Swapping. For original pattern content, the interaction was significant ($F(1,47) = 25.23$, p

$< .0001$; $F(1,31) = 48.82$, $p < .0001$; $\eta_G^2 = .12$) due to the predicted metrical restoration effect.

For swapped-pattern content, the interaction was also significant ($F(1,47) = 10.27$, $p = .002$;

$F(1,31) = 8.10$, $p = .008$; $\eta_G^2 = .03$), but here the interaction resulted from a *negative* restoration

effect. This means that the restoration effect was reversed with respect to the melodies' identity.

This is consistent with an active role for rhythmic patterns in listeners' metrical restoration, and

suggests that effects of rhythmic pattern were stronger than effects of overall melodic form.

Was there also an effect of melodic form? This can be addressed by asking whether the

rhythmic pattern effect was equally strong for the original melodies and the swapped melodies—

that is, rescoring a "match" as a match to the rhythmic pattern, not to the melody itself (third bar

in Figure 7). If rhythmic pattern match is the *only* factor dictating metrical restoration, then the

magnitude of the pattern effect should be just as large as the metrical restoration in the unaltered

melodies. However, if melodic form exerts some influence as well, then pattern-match

restoration should be weaker for the changed melodies. This was tested by rescoring the

swapped-pattern melodies' ratings as match to rhythmic pattern content, and then repeating the

ANOVAs described above. The three-way interaction of Melody Type x Exposure Meter x

Probe Meter did reach significance ($F(1,47) = 8.06$, $p = .007$; $F(1,31) = 6.93$, $p = .01$; $\eta_G^2 = .01$),

with a smaller restoration effect when the melodic form did not concur in meter with the

rhythmic pattern (and the original timbre). This suggests that melodic form, in addition to a

melody's component rhythmic patterns, also affects metrical restoration.

Discussion

The current experiment explored whether changing a melody's rhythmic patterns simply decreases melody recognition, or instead actively reshapes metrical perception. If rhythmic patterns actively shape metrical perception, then the restoration effect should significantly reverse when the component rhythmic patterns are associated with the opposite meter. The restoration effect was significantly in the wrong direction with respect to melody identity when rhythmic pattern (and timbre) changed. That is, when listeners heard Melody 1 with Melody 2's rhythmic patterns, they had a detectable preference for Melody 2's meter. This suggests that rhythmic patterns can strongly dictate metrical perception. Nonetheless, restoration considered at the level of the rhythmic pattern is not as strong when the melody's pattern-and-timbre content is swapped. This suggests that melodic form also influences metrical restoration.

Of course, this does not tell us what the critical aspects of melodic form are, or what cluster of musical attributes might constitute it. In the current experiment, melodic form effects cannot be ascribed to gross differences such as pitch range or overall duration, which are fairly similar across melodies. One possibility is that *contour* is a melodic property related to global form. Contour is often described in terms of note-to-note pitch changes, but one can also characterize contour relationships between nonadjacent notes (e.g. Quinn, 1999; see also Morris, 1993). Perhaps broader-scale contour relationships, such as those between tones in strong metrical positions, or longer-duration contour profiles of melodies over several measures—both of which were consistent across changes in rhythmic patterns in the melodies here (see Appendix C)—constitute at least one aspect of melodic form. This was tested in the final experiment.

Experiment 6: melodies share rhythmic patterns but not contours

This experiment tested whether melodic form cues, which appear to operate independently from rhythmic patterns in Experiment 6, may be related to melodic contour. If contour is indeed an important factor in overall melodic form, then contour should play some role in metrical restoration. Thus, this final experiment assesses whether there is a role for contour in dictating metrical restoration, or if rhythmic structure predominates.

Method

Participants. N=48 new participants from the same pool as previous experiments took part.

Table 7. Experiment 6 melodies and contours.

Melody	Rhythmic patterns	Pitch contour
1	4 1 1 3 1 1 1	fall (fall) fall fall
2	4 1 1 3 1 1 1	rise (rise) rise rise
3	1.5 .5 1 1 2 1 1 1 1 2	rise rise fall fall
6	1.5 .5 1 1 2 1 1 1 1 2	rise fall rise fall
4	2 1 1 2 3 1 2	fall (fall) fall fall
5	2 1 1 2 3 1 2	rise (rise) rise rise
7	2 1 1 1 1 1 1 1 1 1 1	0 (0) 0 0 0
8	2 1 1 1 1 1 1 1 1 1 1	rise (rise) rise fall fall

Stimuli. Stimuli were similar to those used in Experiment 3, where pairs of melodies shared a pair of rhythmic patterns. However, in the current experiment, the two rhythmic patterns

might fall in pitch throughout the measure in Melody 1, but rise in pitch throughout the measure in Melody 2 (see Table 7 and Appendix C). Thus, if listeners represent rhythm + contour, then there should be restoration for Melodies 1 and 2 even if their exposure meters conflict. If listeners only represent rhythmic patterns, then there should be massive interference between the melodies, resulting in no metrical restoration.

Procedure. This matched Experiment 3, with either 4 or 8 exposures to each melody during the exposure phase, and 32 test trials.
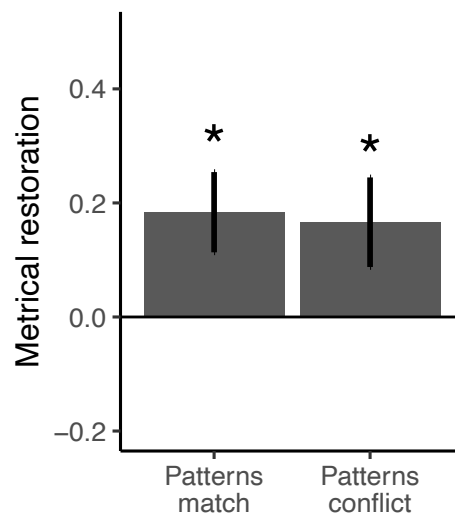

Figure 8. Experiment 6, metrical restoration. * *p* < .05

Results

Figure 8 gives little indication of weaker restoration where different-contour melodies shared the same rhythmic patterns. Metrical fit ratings were analyzed in an ANOVA with Pattern+Contour Conflict (matched meter across instances of the rhythmic pattern, vs. conflicting meters across instances of the rhythmic pattern), Number of Exposures (4, 8), Exposure Meter, and Probe Meter as factors. The effect of Exposure Meter x Probe Meter was

significant ($F1(1,44) = 10.73$, $p = .002$; $F2(1,7) = 10.95$, $p = .01$; $\eta_G^2 = .06$), suggesting that

metrical restoration occurred. The interaction of Pattern+Contour Conflict x Exposure Meter x

Probe Meter did not approach significance ($F1(1,44) = 0.03$, $p = .87$; $F2(1,7) = 0.09$, $p = .77$;

$\eta_G^2 = .00$), implying that metrical restoration was as strong in the conflicting pattern+contour

condition as in the matching pattern+contour condition. Confirming this, each condition

individually showed significant restoration (matching: $F1(1,23) = 4.48$, $p < .05$; $F2(1,7) = 8.04$, p

$= .03$; $\eta_G^2 = .05$; conflicting: $F1(1,23) = 6.82$, $p = .02$; $F2(1,7) = 8.84$, $p = .02$; $\eta_G^2 = .09$). No

other effects approached significance.

Discussion

  This final experiment asked whether melodic contour can distinguish metrical

associations with particular rhythmic patterns, and found a positive answer. Melodies with the

same rhythmic patterns but different contours superimposed on those patterns showed metrical

restoration, even when the meters of those two melodies conflicted (one was in 3/4 time, one was

in 6/8 time). These results suggest that contour is an important ingredient in metrical restoration,

as it appears to prevent interference between rhythmically-similar melodies. These contour

patterns might reflect an aspect of melodic form. They might also reflect that particular rhythm-

plus-contour patterns that extend across a compositional style support meter perception. One

example of such a style-specific rhythm-plus-contour pattern would be the Alberti bass pattern

([down]-up-down-up) seen in accompaniments in classical music. It is open to empirical

assessment how common such figures are. Of course, other factors, such as implied harmony,

remained the same even when rhythmic patterns changed, leaving open whether contour-based

overall form patterns can dictate metrical restoration on their own. That question is left for later investigations.

General Discussion

The current study asked what factors dictate metrical restoration. Findings, summarized in Table 8, indicate strong roles for unique rhythmic patterns, overall melodic form, and contour. Somewhat surprisingly, other factors contributing to the uniqueness of a melody—timbre, rate, or both—did not appear to favor restoration (Experiment 1). Additionally, changing timbre at test appeared to have no effect on restoration (Experiment 4). This is especially surprising in that timbre is known to facilitate recognition memory (Halpern & Müllensiefen, 2007; Radvansky, Fleming, & Simmons, 1995; Radvansky & Potter, 2000), and previous studies have found effects of timbre-specific metrical restoration (Creel, 2012). Nonetheless, diversity of timbres does not impede restoration, as long as rhythmic patterns are consistent within a meter (Experiments 2 and 3). Rhythmic patterns can generate restoration (Experiments 2-4) even if similar rhythmic patterns are associated with the opposite meter (Experiment 2). However, associating a single rhythmic pattern with two conflicting meters, or switching rhythmic patterns at test to a pattern associated with the opposite meter, did cause loss of metrical restoration (Experiments 3, 4, and 5). Still, conflicting rhythmic patterns did not completely reverse restoration, suggesting that other cues—those tied to overall melodic form—hold some sway (Experiment 5). Finally, Experiment 6 suggested that contour, a possible component of overall melodic form, allowed restoration even though melodies' rhythmic patterns mapped to conflicting meters. In summary, rhythmic patterns and more global melodic form cues may be activated in concert to generate a restored percept of meter.

The first major contribution of this study is in showing that rhythmic patterns have a very strong influence on metrical restoration, even when contour is not controlled for. The crux of the effect seems to be that rhythmic patterns, possibly with a contour component, become strongly associated with a particular meter and activate that meter during the listening process. The second major contribution of this study is in showing that overall melodic form cues influence metrical restoration. Overall melodic form cues may include, but are not limited to, coarse-grained melodic contour and implied harmony.

Table 8. Factors found to affect metrical restoration.

Factor	Affects restoration?	Exp.
Timbre uniqueness*	No	1, 4
Rate uniqueness	No	1
Unique rhythmic pattern	*Yes*	**2, 3, 4**
Number of exposures	No? **	2, 3, 6
Rhythmic pattern similarity	No/Weak	2
Unique rhythmic pattern-melody combination	No	3, 4
Rhythmic pattern immediate repetition	No	3
Overall melodic form	*Yes*	**4, 5**
Unique rhythmic pattern-contour combination	*Yes*	**6**

*This conflicts with Creel's (2012) findings. See discussion in text.
**In the limit, exposure would have to have effects, in that *zero* exposures should generate no restoration.

New clarity on factors that may shape metrical restoration

Rhythmic patterns. How exactly do rhythmic patterns shape metrical restoration? According to the current study, a rhythmic pattern simply needs to be uniquely paired with a meter. There is limited "bleed-over" (interfering coactivation) to similar rhythms, though there is strong bleed-over across melodies which share identical rhythms. Further, the rhythm does not

need to cycle repeatedly within a melody for restoration to occur.[2] Figure 9 summarizes the

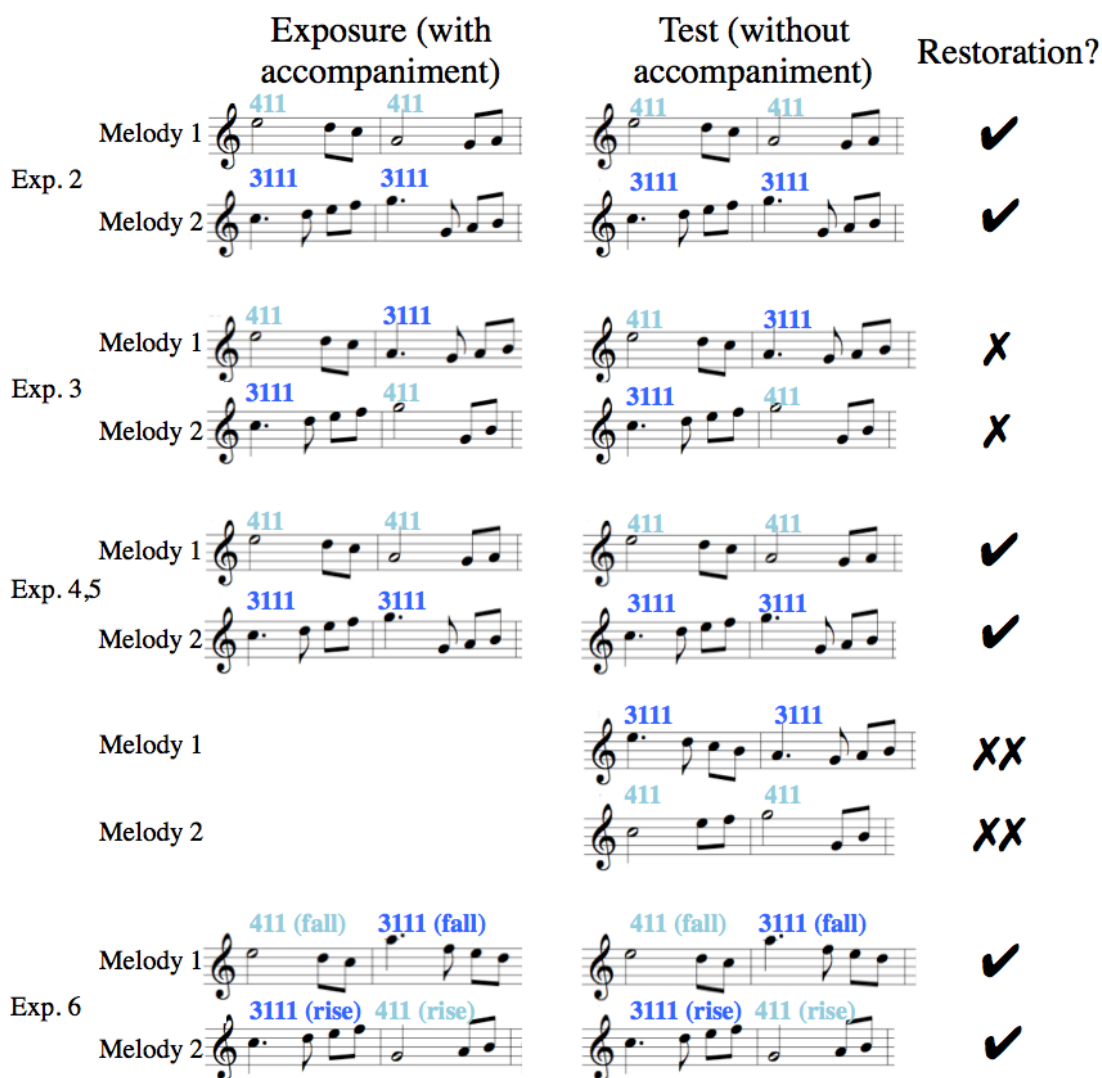distributions of rhythmic patterns in Experiments 2-6, and their effects on metrical restoration.

Figure 9. Summary of effects in Exps. 2-6, showing how rhythmic patterns were distributed across melodies. For each experiment, only "conflict" conditions are depicted—those where melody pairs with similar or shared rhythmic patterns, such as Melodies 1 and 2, mismatched in the meters they occurred with. Numbers indicate rhythmic pattern in each measure. ✔ = showed preference for exposure meter; ✗ = showed no preference; ✗✗ = showed DISpreference for exposure meter

---

[2] Note also that what is talked about here as a rhythmic pattern is more accurately a rhythmic pattern *aligned with metrically strong positions*.

Overall melodic form. Despite strong effects of rhythmic pattern overlap, it cannot be the only factor in metrical restoration. In the current study, overall melodic form influenced meter percepts even when a conflicting-meter rhythm was spliced in. Further, strongly-differing contours distinguished otherwise-identical rhythms from each other. Perhaps most obviously, listeners in the current set of experiments were successfully associating rhythmic patterns with particular meters *even though they have experienced these rhythmic patterns in association with a variety of meters in real life*. Thus, some aspect of the experiment—perhaps the novel melodic forms heard, context-dependent learning in an unfamiliar lab environment, or both—allows learners to associate melodies with meters with relatively little interference.

Melodic form is perhaps also the best explanation for the presence of metrical restoration in previous studies where there was substantial, haphazard rhythmic-pattern overlap across meters (Creel, 2012, 2013). That is, rhythmic patterns were not associated strongly with meters, so rhythmic patterns could not have driven those effects. But why are rhythmic pattern effects so strong relative to melodic form effects, at least in the current study? Perhaps this is because listeners simply had much more exposure to the rhythmic patterns than to overall melodic forms. With greater and greater exposure to melodic forms, perhaps their influence would increase. This might predict that massive exposure to melodies in Experiment 1, increasing melodic form familiarity, would strengthen metrical restoration. On the other hand, rhythmic patterns may have a stronger influence than melodic form because there is less similarity-based interference amongst rhythmic patterns than amongst melodic forms.

Timbre. In the current study, timbre specificity appeared to have no influence on metrical restoration. However, Creel (2012, Experiments 1 and 2) showed effects of timbre specificity, such that changing a melody's timbre to a timbre associated with the opposite meter blocked

metrical restoration. Why are the current results so different from Creel's (2012) findings — that is, when does timbre have an effect? One possibility is that Creel's (2012) Experiments 1 and 2 used melodies without associations between rhythmic patterns and particular meters, while the current experiments did have strong rhythmic pattern-meter associations. Perhaps timbre only shows its influence when rhythmic pattern associations are weak.

Another possibility is that timbre may affect processing not by keeping melodies *separate* from each other, but by *allowing coactivation of* distinct melodies. That is, if melodies are similar in timbre, those traces will tend to get coactivated during metrical restoration and support that restoration. This fits with the timbre-specificity in the two-timbre experiments from Creel (2012). However, if *no* melodies are similar in timbre (and no rhythmic patterns support a particular meter), no other traces will be coactivated and metrical restoration will be weak. This fits with the timbre-diverse conditions in the current Experiment 1, which showed no evidence of metrical restoration. It is also possible that with extensive exposure, timbre uniqueness might become a good cue for metrical restoration, but that remains to be tested.

Further questions

Do metrical restoration effects generalize to real-world listening? The account here is that restoration of missing information is a general phenomenon in music perception: listening is guided by large-scale, lifetime distributional pattern learning of musical exemplars (Goldinger, 1998; Hintzman, 1986; Pierrehumbert, 2001). As distributions accumulate over a listener's lifespan, patterns emerge that become activated when hearing new music and shape meter perception. As a signal is heard, it activates stored representations gradiently, with representations more similar to the input becoming more active and generating a composite

"echo" that influences the percept (Hintzman, 1986). In the experimental settings presented here, hearing an isolated melody should activate previous hearings of that melody along with its accompaniment, and will also activate hearings of other melodies to the extent that they are similar. If no other melodies are similar enough to receive strong activation, then the "echo" of that melody alone will dominate, filling in the previous meter most strongly. However, if other melodies are highly similar, then the echo will contain elements of multiple melodies, and if those melodies vary in meter, the metrical information in the composite echo will cancel out to some extent.

The working assumption throughout this paper has been that the same memory processes generating metrical restoration effects observed in the lab are a microcosm of more interesting real-world listening situations, including musical style-specific expectations, and normal listening conditions. This fits with many laboratory studies, such as the statistical language learning studies by Saffran and colleagues (Saffran, Aslin, & Newport, 1996; Saffran, Newport, & Aslin, 1996), which assume that the cognitive mechanisms that allow learning of a few words in a novel, artificial language in a lab setting will scale up to learning full human languages of tens of thousands of words.

What, then, happens in real-world listening? While some of the coactivated musical exemplars in the current study might include melodies from lifelong learning (enculturation), it is likely that exemplars stored in the lab setting have a recency advantage (see Pierrehumbert, 2001) and will dominate. But considering music listening in the real world, the relevant space of music is much, much larger, and exposure to particular musical pieces is likely greater. This might have three effects. The first is a smoothing of the echo toward something more like an abstracted representation, because a greater diversity of exemplars contribute to it. Second is a

stronger echo due to greater numbers of exemplars (relative to lab-acquired music) from frequent replays of recorded music and redundancy within a single piece (repeated verses, repeated themes). Third, it is possible that various musical attributes shift in their relative importance in meter processing after massive amounts of exposure.

The experiments might seem to suggest that rhythmic patterns shared across meters would overlap so much as to generate catastrophic interference in real-world music processing, leading to poor meter perception. In natural listening experiences, such overlap is presumably rampant, yet most listeners readily perceive differences amongst meters in music of their culture (e.g. Drake & El Heni, 2003; Iversen & Patel, 2008), suggesting that such interference is not problematic. As noted above, much lengthier exposure to familiar music might lead to stronger metrical representations, stronger melodic-form representations, or both. Additionally, melodic cues to meter that are commonly found in real music were deliberately minimized in the current study, but a variety of cues such as pitch and timing accents (Ellis & Jones, 2009; Hannon, Snyder, Eerola, & Krumhansl, 2004) and, of course, meters in accompanying musical lines are known to facilitate meter perception.

Is memory-based metrical processing obligatory? If one takes the view that unambiguous meter cues are typically present during natural listening, is memory-based metrical processing really necessary? A better question may be, is memory reactivation *obligatory*, as the exemplar account outlined here suggests? Research on cross-cultural meter perception suggests that the answer is yes. Recall that Hannon and Trehub (2005a, 2005b; see also Hannon et al., 2012; Kalender et al., 2013) have found that Western listeners, who have had little to no exposure to complex meters, have difficulty processing those complex meters. Essens and Povel (1985) and Snyder, Hannon, Large, and Christiansen (2006) have found that Western listeners partly

regularize these complex patterns to more familiar Western ones—even when hearing the melodies themselves, which presumably contain cues to metrical prominence (Snyder et al., 2006). An exemplar account of these phenomena suggests that listeners experience interference from their native musical representations. Thus, without a large collection of memory traces via massive cultural exposure to support meter perception, listeners are quite poor at using any of the musical surface cues that are presumably present even in complex-meter music.

It is possible that adult-like perception of simple meters is also bolstered by lengthy cultural exposure. While numerous real-world musical signals doubtless contain unambiguous metrical information, many noise sources may partially obscure real-world musical signals: reverberation, expressive timing variations, and actual noise sources (air conditioners, traffic, the babble of surrounding voices). If real-world listening *is* noisy, then top-down pattern completion would be advantageous for music recognition, including meter recognition. Thus, there is an argument to be made that top-down information like an exemplar echo is often, if not always, useful.

Implications for models of musical meter perception. While current results are generally consistent with an exemplar account of memory, explicit model findings are absent. To the author's knowledge, no current model of meter detection or metrical entrainment can account for the results reported here (for example, Large, Herrera, & Velasco, 2015; Tomic & Janata, 2008). This is because those models all use bottom-up information—the immediate musical input— rather than a combination of bottom-up and top-down information. It is possible that distributional-learning models such as Pearce and Wiggins' (2012) IDyOM model might be adapted to account for some of these effects, though IDyOM is not explicitly an exemplar model and currently is designed to account for musical pitch processing rather than timing. Nonetheless,

the current study suggests that current models should be modified or elaborated to account for effects of detailed musical memories on processing, particularly if these models wish to account for differential patterns in different musical styles.

Conclusion

The study explored influences on metrical restoration, a phenomenon in which listeners who hear melodies with ambiguous meter report preferences for meters that are influenced by specific previous listening experience in the lab. The main components guiding metrical restoration appear to be rhythmic pattern-meter associations, and some aspect of overall melodic form. Overall melodic form may be related to patterns of contour across the melody. Metrical restoration-like effects may underlie meter perception in normal listening situations as well, a topic for future experimental investigation and modeling efforts.

References

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

Barnes, R., & Jones, M. R. (2000). Expectancy, attention, and time. *Cognitive Psychology*, *41*(3), 254–311.

Bhatara, A., Boll-Avetisyan, N., Agus, T., Höhle, B., & Nazzi, T. (2016). Language experience affects grouping of musical instrument sounds. *Cognitive Science*, *40*, 1816–1830. http://doi.org/10.1111/cogs.12300

Bhatara, A., Boll-Avetisyan, N., Unger, A., Nazzi, T., & Höhle, B. (2013). Native language affects rhythmic grouping of speech. *Journal of the Acoustical Society of America*, *134*(5), 3828–43. doi:10.1121/1.4823848

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Casey, M. A., Veltkamp, R., Goto, M., Leman, M., Rhodes, C., & Slaney, M. (2008). Content-Based Music Information Retrieval: Current Directions and Future Challenges. In A. Hanjalic, R. Lienhart, W.-Y. Ma, & J. R. Smith (Eds.), *Proceedings of the IEEE* (Vol. 96, pp. 668–696).

Creel, S. C. (2011). Specific previous experience affects perception of harmony and meter. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1512–1526.

Creel, S. C. (2012). Similarity-based restoration of metrical information: Different listening experiences result in different perceptual inferences. *Cognitive Psychology*, *65*(2), 321–51.

Creel, S. C. (2013). Dimensions of specificity in musical memory: Evidence from metrical

restoration. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual Conference of the Cognitive Science Society* (pp. 2106–2111). Austin, TX: Cognitive Science Society.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, *25*(4), 385–400.

Dilley, L. C., Mattys, S. L., & Vinke, L. (2010). Potent prosody: Comparing the effects of distal prosody, proximal prosody, and semantic context on word segmentation. *Journal of Memory and Language*, *63*(3), 274–294.

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.

Drake, C., & El Heni, J. Ben. (2003). Synchronizing with music: intercultural differences. *Annals of the New York Academy of Sciences*, *999*, 429–437.

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), 264–80.

Essens, P. J., & Povel, D.-J. (1985). Metrical and nonmetrical representations of temporal patterns. *Perception & Psychophysics*, *37*(1), 1–7.

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

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), 956–74.

Hannon, E. E., Soley, G., & Ullal, S. (2012). Familiarity overrides complexity in rhythm

    perception: a cross-cultural comparison of American and Turkish listeners. *Journal of*

    Experimental Psychology: Human Perception and Performance, *38*(3), 543–8.

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood.

    Psychological Science, *16*(1), 48–55.

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more

    readily than adults. *Proceedings of the National Academy of Sciences*, *102*(35), 12639–

    12643.

Iversen, J. R., & Patel, A. D. (2008). The Beat Alignment Test (BAT): Surveying beat

    processing abilities in the general population. In M. Adachi (Ed.), *Proceedings of the*

    10th International Conference on Music Perception and Cognition (pp. 2–5). Sapporo,

    Japan: Causal Productions.

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), 2263–71.

Iverson, P., & Krumhansl, C. L. (1993). Isolating the dynamic attributes of musical timbre.

    Journal of the Acoustical Society of America, *94*(5), 2595–603. Retrieved from

    http://www.ncbi.nlm.nih.gov/pubmed/8270737

Jones, M. R., Moynihan, H., Mackenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-

    driven attending in dynamic arrays. *Psychological Science*, *13*, 313–319.

Kalender, B., Trehub, S. E., & Schellenberg, E. G. (2013). Cross-cultural differences in meter

    perception. *Psychological Research*, *77*(2), 196–203.

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

Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in

musical rhythm. *Frontiers in Systems Neuroscience*, *9*(November), 1–14.

doi:10.3389/fnsys.2015.00159

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: MIT Press.

London, J. (2002). Cognitive constraints on metric systems: Some observations and hypotheses.

Music Perception, *19*(4), 529–550.

Morris, R. D. (1993). New directions in the theory and analysis of musical contour. *Music

Theory Spectrum*, *15*(2), 205–228.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect

size for some common research designs. *Psychological Methods*, *8*(4), 434–47.

Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical

rhythms. *Music Perception, 11*(4), 409–464.

Pearce, M. T., & Wiggins, G. A. (2012). Auditory expectation: the information dynamics of

music perception and cognition. *Topics in Cognitive Science*, *4*, 625–652.

doi:10.1111/j.1756-8765.2012.01214.x

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers

into movies. *Spatial Vision*, *10*(4), 437–442.

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

411–440.

Quinn, I. (1999). The Combinatorial Model of Pitch Contour. *Music Perception*, *16*(4), 439–456.

Radvansky, G. A., Fleming, K. J., & Simmons, J. A. (1995). Timbre reliance in nonmusicians'

and musicians' memory for melodies. *Music Perception*, *13*(2), 127–140.

Radvansky, G. A., & Potter, J. K. (2000). Source cuing: memory for melodies. *Memory & Cognition*, *28*(5), 693–9.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word Segmentation: The Role of Distributional Cues. *Journal of Memory and Language*, *35*(4), 606–621.

Samuel, A. G. (1981). Phonemic restoration: insights from a new methodology. *Journal of Experimental Psychology: General*, *110*(4), 474–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2939193

Scheirer, E. D. (1998). Tempo and beat analysis of acoustic musical signals. *Journal of the Acoustical Society of America*, *103*(1), 588–601.

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

Schenker, H. (1979). *Free composition* (E. Ouster, Ed. and Trans.). New York: Longman. (Original work published 1935).

Snyder, J. S., Hannon, E. E., Large, E. W., & Christiansen, M. H. (2006). Synchronization and continuation tapping to complex meters. *Music Perception*, *24*(2), 135–148.

Temperley, D., & Bartlette, C. (2002). Parallelism as a factor in metrical analysis. *Music Perception*, *20*(2), 117–149.

Tomic, S. T., & Janata, P. (2008). Beyond the beat: Modeling metric structure in music. *Journal of the Acoustical Society of America, 124*(6), 4024–4041. doi:10.1121/1.3006382

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., & Werker, J. F.

(2010). The development of perceptual grouping biases in infancy: a Japanese-English

cross-linguistic study. *Cognition*, *115*(2), 356–61.

Appendix A: Music experience and metrical restoration

Table A1. Effects of music experience on metrical restoration.

Exp.	N	Gender	Age	Music experience[c]	Correlation with metrical restoration	Music experience differences between groups
						Diverse-timbre > single timbre, p
1	72[a]	52 F, 17 M[b]	21.0 (2.4)	6.1 (5.7)	0.217+	= .009
2	96	71 F, 25 M	20.8 (2.1)	6.1 (6.2)	0.222*	.
3	48	32 F, 16 M	21.3 (1.8)	4.6 (5.5)	0.530***,d	ns
4	48	42 F, 6 M	20.8 (2.7)	8.1 (6.2)	0.308*	ns
4ctrl	48	33 F, 15 M	20.1 (1.8)	6.7 (6)	0.229	.
5	48	38 F, 10 M	20.3 (1.6)	5.9 (5.7)	-0.052	.
6	48[a]	35 F, 12 M[b]	20.0 (1.5)	6.4 (5.6)	0.134	ns

Note. Music experience was quantified as the number of years during which the participant stated they had played a musical instrument or sung.
[a] Two participants in Experiment 1, and one in Experiment 6, failed to complete background questionnaires and are not included in the analyses here.
[b] One participant declined to provide gender information.
[c] While means were fairly high, many participants reported 0 years of music experience (Exp. 1: 32%; Exp. 2: 36%; Exp. 3: 48%; Exp. 4: 19%; Exp. 4 control: 38%; Exp. 5: 35%; Exp. 6: 34%).
[d] The top four metrical restoration scores were in the Match condition. All four participants had 12 or more years music experience. Removing them dropped the correlation to .21, but the Conflict effect was still significant, $p = .009$.

Appendix B: Control experiment to test metrical biases of melodic stimuli

Method. New participants (N=24) heard the twelve melodies from Experiment 1 and the

eight melodies from Experiment 2 and rated the goodness-of-fit of 3/4 probes and 6/8 probes. All

melodies were presented without accompaniment, in piano timbre. ANOVAs were computed

with Probe Meter (3/4, 6/8) and Experiment (Exp. 1, Exp. 2) as factors.

General results. There was a mild preference for 3/4 probes (effect of Probe Meter:

$F1(1,22) = 3.72, p = .07; F2(1,18) = 3.41, p = .08; \eta_G^2 = .09$), but neither the effect of

Experiment nor the interaction approached significance (all Fs < 1, ps ≥ .50; $\eta_G^2 = .00$). This

suggests that the mild 3/4 preference was consistent across experiments, and that there was no

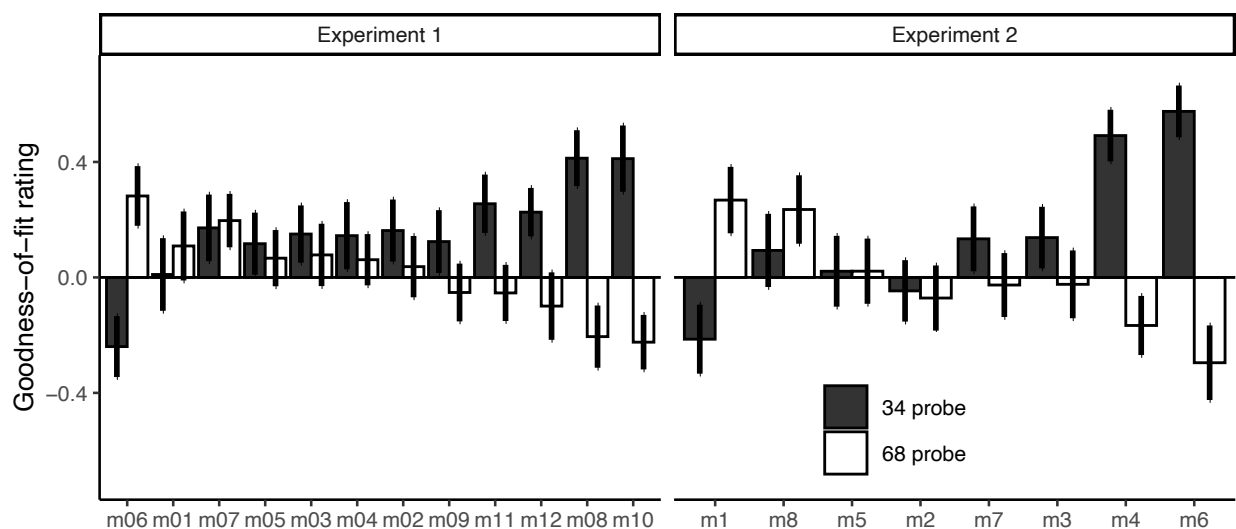difference in baseline 3/4 or 6/8 biases between experiments.

Figure B1. Control experiment, metrical probe ratings, ± standard errors.

Individual melodies had different metrical biases, but biases did not block metrical

restoration effects. There were differences among individual melodies in their degree of 3/4 vs.

6/8 preference (Figure B1). For each melody in the control experiment, I calculated a 3/4 bias

score (3/4 probe rating for that melody minus 6/8 probe rating for that melody). I also calculated

overall 3/4 bias for each melody in Experiments 1 and 2, without respect to whether each probe

was "correct" (matched memory) or not. To assess whether 3/4 bias affected metrical probe

ratings, I computed correlations between 3/4 bias in the control experiment and 3/4 bias in the
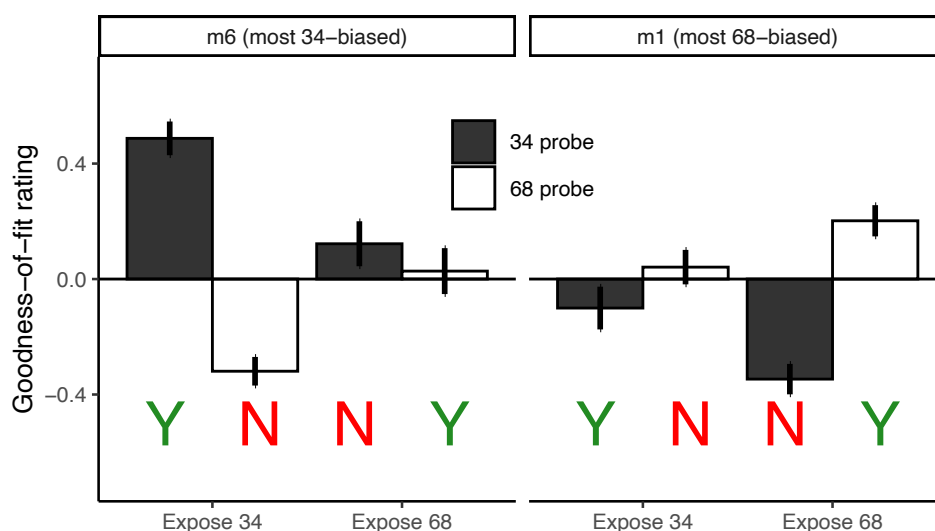

Figure B2. Experiment 2, probe ratings for the most 6/8-biased and the most 3/4-biased melody
according to the control study. Experiment 2 ratings for the 6/8-biased melody were overall
higher for 6/8, and the reverse was true for the 3/4 biased melody. Nonetheless, exposure effects
were evident for both melodies in that exposure-matched probe ratings (Y's) were higher than
exposure-mismatched ratings (N's).

main experiments. These bias scores correlated positively in each experiment (Exp. 1: $r(10)$

$= .56, p = .06$; Exp. 2: $r(6) = .78, p = .02$). This suggests that *baseline* metrical biases of melodies

used in Experiments 1 and 2 affected probe ratings, consistent with previous findings that timing

and pitch cues in melodies exert effects on meter perception (Ellis & Jones, 2009; Hannon et al.,

2004). However, because melody-to-accompanying-meter assignments were counterbalanced

across participants (for each person who heard Melody 1 in 3/4, another heard it in 6/8), these

effects are orthogonal to exposure effects. Rather, it suggests that the learning effects seen in

Experiments 1 and 2 combine with cues in the melodies themselves (Figure B2).

Individual melodies' bias to either meter did not differ between experiments. Is it possible that there were weaker exposure effects in Experiment 1 because that experiment's melodies were more biased toward one or the other meter, blocking restoration effects? To assess this, I calculated the *absolute value* of the bias score (| 34 rating – 68 rating |) for each melody in the control experiment. This reflects how large the overall preference is in either direction (for either 3/4 or 6/8). Absolute bias scores for Experiment 1 melodies were compared to scores for Experiment 2 melodies in a *t*-test. This test missed significance (t(11.6) = 0.45, p = .66, d = .21; the non-integer degrees of freedom represents a correction for unequal variances), suggesting that individual Experiment 1 melodies were no more meter-biased than Experiment 2 melodies.

Variability of individual subjects' ratings of a melody differed across experiments, but this did not appear to drive exposure effects. While *mean* ratings of melodies did not differ between experiments, the *variability* of ratings did differ across experiments. Specifically, 3/4 bias scores were more variable across participants (higher SD) for Experiment 2 melodies (average of standard deviations: .86 ± .08) than for Experiment 1 melodies (.76 ± .08; t(14.7) = 2.83, p = .01, d =1.30). It is not immediately clear that this should affect learnability of exposure meters, but one might conjecture that more-variably-perceived melodies (as in Experiment 2) might be more learnable than less-variably-perceived melodies (as in Experiment 1). If so, then individual melodies' malleability should correlate positively with metrical restoration *within* each experiment. To test this, I computed correlations between the *standard deviation of 3/4 bias* from the control study with *size of metrical restoration score* for each melody within each Experiment. However, these values were nonsignificantly correlated in the wrong direction (Experiment 1: r(10) = -.09, p = .78; Experiment 2: r(6) = -.31, p = .46). This suggests that higher variability in judgments of a melody's meter has little influence on exposure effects.

Appendix C
Example stimuli from Experiments

Figure C1. Experiment 1, six example melodies. Measure-long rhythmic patterns are notated numerically above each measure.

Figure C2. Experiment 2, first half of each of the eight melodies. Rhythmic patterns are marked numerically over each measure. For present purposes, dotted half notes were not considered to be shared patterns.

Melody 1

Melody 2

Figure C3. Experiment 3, example pair of melodies with shared rhythmic patterns (4 1 1, 3 1 1
1), first 8 measures of each melody. In the *matching* condition, both would be heard with 3/4
accompaniments (or both 6/8). In the *conflicting* condition, the first would be heard in 3/4 and
the second in 6/8 (or vice versa).

Melody 1, pattern 1 (4 1 1)

Melody 1, pattern 2 (3 1 1 1)

Melody 2, pattern 1 (4 1 1)

Melody 2, pattern 2 (3 1 1 1)

Figure C4. Experiment 4, example melodies with changed rhythmic patterns (first 8 measures of
each melody).

Melody 1, falling pitch

Melody 2, rising pitch

Figure C5. Experiment 6, melodies sharing rhythmic patterns but differing in pitch contours.

Appendix D:
Goodness-of-fit ratings in each experiment split by exposure meter and probe meter.

Table D1. Experiment 1, means (SDs) of goodness-of-fit ratings by condition, exposure meter, and probe meter.

Timbres	Rates	Exposure meter	34 probe	68 probe
Single (1)	Single (1)	34	**0.160 (0.394)**	-0.253 (0.363)
		68	0.061 (0.353)	**-0.075 (0.388)**
	Diverse (3)	34	**0.222 (0.284)**	-0.101 (0.171)
		68	0.118 (0.280)	**0.053 (0.257)**
Diverse (12)	Single (1)	34	**0.184 (0.374)**	0.005 (0.378)
		68	0.102 (0.392)	**0.102 (0.370)**
	Diverse (3)	34	**0.236 (0.288)**	-0.005 (0.238)
		68	0.222 (0.219)	**0.025 (0.201)**

Note. Throughout, entries in **bold** indicate a match between exposure and probe meters.

Table D2. Experiment 2, goodness-of-fit ratings by condition, exposure meter, and probe meter.

Number Exposures	Exposure meter	34 probe	68 probe
2	34	**0.107 (0.365)**	-0.033 (0.265)
	68	-0.064 (0.354)	**0.121 (0.238)**
4	34	**0.132 (0.302)**	-0.042 (0.271)
	68	-0.042 (0.280)	**0.080 (0.291)**
8	34	**0.235 (0.354)**	-0.136 (0.253)
	68	-0.138 (0.254)	**0.173 (0.295)**
16	34	**0.116 (0.402)**	-0.098 (0.315)
	68	-0.126 (0.397)	**0.164 (0.403)**

Table D3. Experiment 3, goodness-of-fit ratings by condition, exposure meter, and probe meter.

Tested rhythm	Number Exposures	Exposure meter	34 probe	68 probe
Match	4	34	**0.002 (0.415)**	-0.077 (0.419)
		68	-0.115 (0.337)	**0.180 (0.263)**
	8	34	**0.326 (0.281)**	-0.301 (0.228)
		68	-0.079 (0.325)	**0.237 (0.305)**
Conflict	4	34	**-0.018 (0.310)**	-0.073 (0.297)
		68	-0.035 (0.306)	**-0.098 (0.277)**
	8	34	**0.057 (0.461)**	-0.101 (0.368)
		68	-0.037 (0.462)	**-0.052 (0.398)**

Table D4. Experiment 4, goodness-of-fit ratings by condition, exposure meter, and probe meter.

Rhythmic pattern	Tested timbre	Tested rhythms	Exposure meter	34 probe	68 probe
Match	Original	Original	34	**0.153 (0.483)**	-0.111 (0.393)
			68	-0.003 (0.429)	**0.149 (0.274)**
		Swapped	34	**0.216 (0.407)**	-0.159 (0.351)
			68	-0.004 (0.397)	**0.051 (0.280)**
	Swapped	Original	34	**0.425 (0.439)**	-0.069 (0.381)
			68	0.004 (0.343)	**0.093 (0.312)**
		Swapped	34	**0.336 (0.433)**	-0.208 (0.368)
			68	-0.026 (0.378)	**0.054 (0.368)**
Conflict	Original	Original	34	**0.160 (0.315)**	-0.090 (0.275)
			68	-0.020 (0.323)	**-0.017 (0.390)**
		Swapped	34	**-0.056 (0.339)**	-0.032 (0.251)
			68	-0.013 (0.291)	**-0.065 (0.241)**
	Swapped	Original	34	**0.153 (0.411)**	-0.014 (0.312)
			68	-0.172 (0.367)	**0.024 (0.440)**
		Swapped	34	**-0.068 (0.307)**	0.064 (0.377)
			68	-0.030 (0.390)	**-0.009 (0.369)**

Table D5. Experiment 5, goodness-of-fit ratings by condition, exposure meter, and probe meter.

Tested rhythm	Exposure meter	34 probe	68 probe
Original	34	**0.252 (0.251)**	-0.188 (0.252)
	68	-0.010 (0.226)	**0.094 (0.294)**
Swapped	34	**-0.015 (0.256)**	0.015 (0.284)
	68	0.123 (0.242)	**-0.089 (0.256)**

Table D6. Experiment 6, goodness-of-fit ratings by condition, exposure meter, and probe meter.

Rhythms	Exposure meter	34 probe	68 probe
Match	34	**0.178 (0.423)**	-0.144 (0.352)
	68	-0.006 (0.361)	**0.004 (0.394)**
Conflict	34	**0.367 (0.324)**	-0.241 (0.300)
	68	0.186 (0.354)	**-0.055 (0.374)**