

# Current Biology

## Integer Ratio Priors on Musical Rhythm Revealed Cross-culturally by Iterated Reproduction

### Highlights

- Perceptual priors on rhythm are revealed by iterated reproduction of random sequences
- Priors in US listeners show peaks at integer ratio rhythms prevalent in Western music
- Priors in native Amazonians are distinct but also favor some integer ratio rhythms
- Musicians have similar priors as non-musicians, implicating passive exposure to music

### Authors

Nori Jacoby, Josh H. McDermott

### Correspondence

[kj2413@columbia.edu](mailto:kj2413@columbia.edu)

### In Brief

Jacoby and McDermott introduce a method to estimate priors on rhythm in human listeners. Both US participants and members of an Amazonian society exhibited priors that give high probability to rhythms whose time intervals form integer ratios. However, the priors otherwise differed across cultures, suggesting strong influences of musical experience.

# Integer Ratio Priors on Musical Rhythm Revealed Cross-culturally by Iterated Reproduction

Nori Jacoby<sup>1,2,3,\*</sup> and Josh H. McDermott<sup>1</sup>

<sup>1</sup>Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>2</sup>The Center for Science and Society, Columbia University, New York, NY 10027, USA

<sup>3</sup>Lead Contact

\*Correspondence: [kj2413@columbia.edu](mailto:kj2413@columbia.edu)

<http://dx.doi.org/10.1016/j.cub.2016.12.031>

## SUMMARY

Probability distributions over external states (priors) are essential to the interpretation of sensory signals. Priors for cultural artifacts such as music and language remain largely uncharacterized, but likely constrain cultural transmission, because only those signals with high probability under the prior can be reliably reproduced and communicated. We developed a method to estimate priors for simple rhythms via iterated reproduction of random temporal sequences. Listeners were asked to reproduce random “seed” rhythms; their reproductions were fed back as the stimulus and over time became dominated by internal biases, such that the prior could be estimated by applying the procedure multiple times. We validated that the measured prior was consistent across the modality of reproduction and that it correctly predicted perceptual discrimination. We then measured listeners’ priors over the entire space of two- and three-interval rhythms. Priors in US participants showed peaks at rhythms with simple integer ratios and were similar for musicians and non-musicians. An analogous procedure produced qualitatively different results for spoken phrases, indicating some specificity to music. Priors measured in members of a native Amazonian society were distinct from those in US participants but also featured integer ratio peaks. The results do not preclude biological constraints favoring integer ratios, but they suggest that priors on musical rhythm are substantially modulated by experience and may simply reflect the empirical distribution of rhythm that listeners encounter. The proposed method can efficiently map out a high-resolution view of biases that shape transmission and stability of simple reproducible patterns within a culture.

## INTRODUCTION

Perception operates on noisy data. Optimal inference thus requires integrating incoming sensory information with prior knowledge. In many areas of perception and cognition, humans

appear to combine current observations with internal beliefs about the environment (the prior) in a process approximating statistical inference [1]. Priors have now been characterized in vision [2–6], time perception [7], audition [8], and motor control [9]. The effect of a prior on an observer’s estimate of the world is usually to bias the estimate toward states that have high probability under the prior [4]. As a result, priors for domains in which signals are reproduced by humans, such as music and speech, directly constrain cultural transmission: only signals with high probability under the prior are reliably reproduced and communicated [10]. However, priors for cultural artifacts have only recently begun to be characterized for language [6] and speech [11] and are unexplored for music.

Priors on music are presumably essential for extracting musical structure from sound [12]. Because of their likely role in cultural transmission, it also seemed plausible that music priors could provide insight into the origins of salient characteristics of music. Although music is highly variable across societies, it is usually defined in part by the structured arrangement of events in time to form rhythm. Particular types of rhythms—those comprised of time intervals related by ratios of small integers (“simple” integer ratios)—are believed to be common, if not universal [13]. What underlies their prevalence? One possibility is that integer ratio rhythms are more easily produced given our motor constraints [14]. Alternatively, representational constraints could affect what can be accurately perceived [15, 16]. To explore this latter possibility, we examined whether integer ratios are reflected in the prior that constrains human perception and transmission of rhythms and whether this prior varies with musical experience and cultural exposure.

Previous probabilistic models of musical structure and/or perception assume priors [17, 18], but to our knowledge priors on musical structure have never been measured, in part because of methodological limitations. Evidence for perceptual priors is traditionally obtained via biases in discrimination or categorization judgments [4]. For music, such tasks are typically difficult, and most results have been obtained with expert musicians [18–20]. Moreover, discrimination experiments are limited in dimensionality because they require large numbers of judgments to establish discriminability at each point in a stimulus space. Rhythm discrimination experiments are thus typically designed to test particular hypotheses derived from Western music theory and cannot probe the large space of possible priors, particularly when considering non-Western cultures for which scientists may not have hypotheses. Experiments in which participants tap to a beat reveal biases that may reflect perception [16, 21, 22] but

have not explicitly characterized a prior. Other approaches to studying rhythm representation require judgments using Western musical notation [18, 20] and are also unsuitable for non-experts or non-Western listeners.

We developed a method to measure perceptual priors on rhythm that is efficient and implicit. Our paradigm is an extension of the game of “broken telephone” previously used to reveal priors in other domains [6, 10, 23]. Listeners are presented with a random “seed” sequence (clicks or other sounds separated by random time intervals) and reproduce it by tapping. In practice, the reproduction is usually performed synchronous with the seed while it is repeatedly presented (because this is easiest for participants), but it can also be performed after seed presentation. The listener’s reproduction is then substituted for the seed, and the process is iterated. If the listener’s response depends on the Bayesian integration of the currently heard rhythm with a prior, and if production biases are small, this process converges to a sample from the prior for the domain, and the prior can be estimated by applying the procedure multiple times [24].

Our method is hypothesis-neutral (allowing any possible pattern of results to be detected), can be applied irrespective of the participant’s musical or cultural background, and estimates the prior over the entire space of simple rhythms. Although related methods are sometimes used to simulate cultural evolution [25, 26], here our goal is strictly to measure the prior within groups of listeners with a common cultural background, in whom the prior might promote cultural stability.

## RESULTS

Our paper is divided into three parts. First, we show the results of applying our method to simple rhythms, to illustrate the approach. Second, we present a series of control experiments to substantiate the claim that the results reflect a perceptual prior, showing that they do not reflect productive biases, that synchronization and replication from memory yield similar results, and that the measured distribution predicts discriminability. Third, we use the method to test the effect of culture and musical experience on the perceptual prior for rhythm.

### Experiment 1: Priors on Rhythm via Iterated Reproduction

Three-interval rhythms are those formed by a repeating sequence of three time intervals. The intervals can be described in terms of their ratios, which for Western musical rhythms can be approximated with integers (Figure 1A) but which in general could be arbitrary non-integer proportions (Figure 1B). The three intervals are produced by four events (e.g., clicks). When the rhythm is presented repeatedly, the fourth event forms the end of the last interval and the beginning of the first interval (Figure 1C).

Western musical notation can only be used to describe rhythms with simple integer ratios (Figure 1D), and so to describe rhythms more generally we will use the ratios between the intervals, plotted in a “rhythm space” [20]. Three-interval rhythms can be represented in three dimensions corresponding to the three intervals (Figure 1E). In this paper, we focus on rhythms whose overall duration (the sum of the intervals) is fixed (typically 2,000 ms, but 1,000 ms for experiments 3, S1, and S2). Because

the overall duration is fixed, the intervals can be presented on a two-dimensional triangular simplex (Figure 1F) known as a “chronotopological map” [20, 27]; for simplicity, here we refer to it as the “rhythm simplex.” Every three-interval rhythm with a fixed duration can be uniquely described by a point on the simplex.

In order to avoid presenting rhythms with intervals too short to reproduce [28], we constrained intervals in the seeds to a smaller simplex (inner triangle in Figure 1E) such that all intervals were larger than 15% of the total duration. All results figures display responses on this inner simplex.

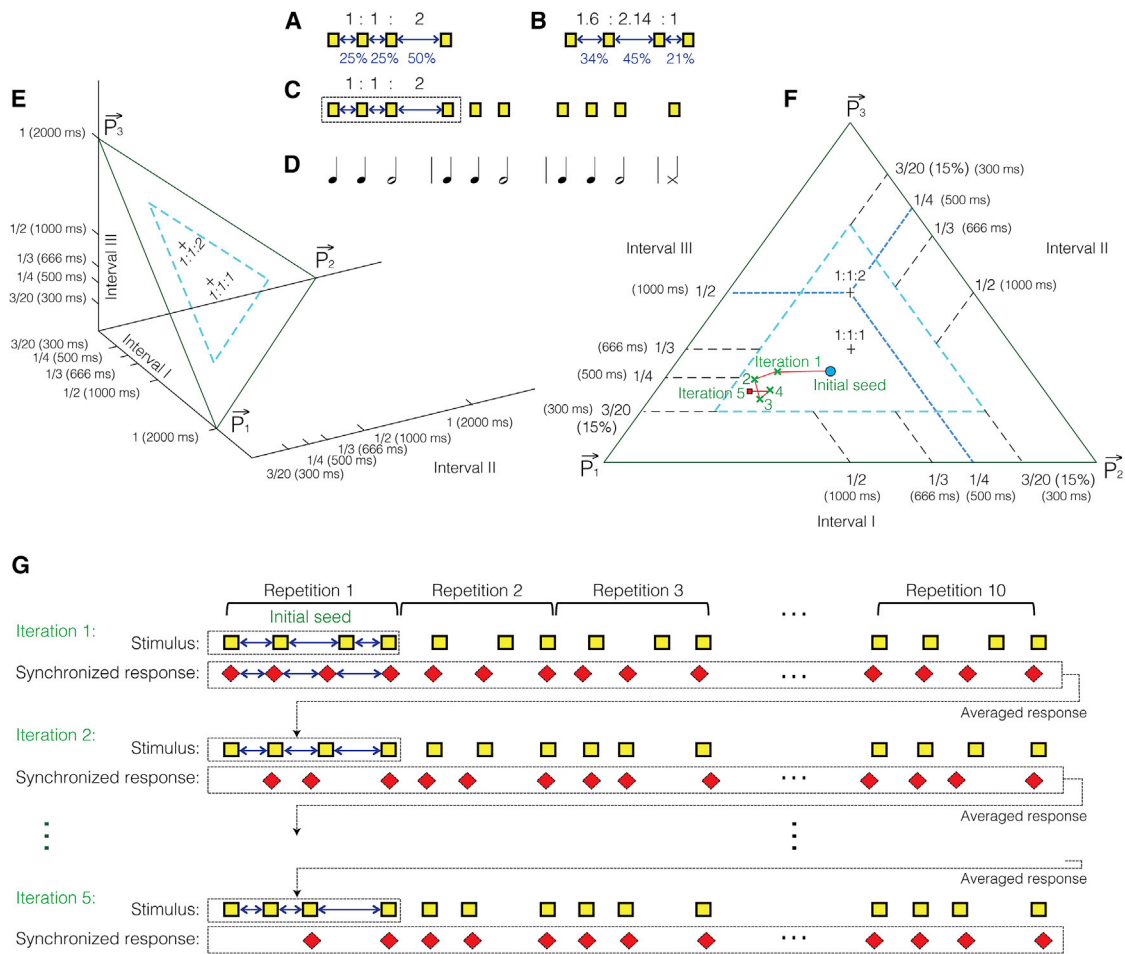
Each trial started with a random seed sampled uniformly from the rhythm simplex (Figure 1F). Participants were presented with ten repetitions of the seed and instructed to synchronize their taps to the repeated rhythm. The inter-tap intervals were averaged across the ten repetitions, and the resulting mean reproduction was then substituted for the seed (Figure 1G). The process was iterated five times.

Figures 2A and 2B show aggregated responses from all participants ( $n = 26$ ; 20–25 seeds per individual; 609 total distinct seeds; participants were a mix of musicians and non-musicians) shown as points on the rhythm simplex for each iteration. The distribution of responses evolves over time and appears to converge to a multi-modal distribution. Copying accuracy (the distance between stimulus and reproduction) improved across iterations (mean stimulus-reproduction distance for iterations 1–5 was 61, 45, 38, 34, and 32 ms, respectively) but was not significantly different for the last two iterations ( $t(25) = -1.85$ ,  $p = 0.08$ ), suggesting that convergence is reached or nearly reached by the fifth iteration. To estimate the continuous distribution underlying the responses, we applied kernel density estimation to each iteration’s data (shown in Figure 2C for the fifth and final iteration).

As expected, the initial distribution of seeds was not significantly different from uniform (measured by the Jensen-Shannon divergence;  $JSD = 0.039$ ,  $p = 0.82$ ). However, the distributions in all subsequent iterations differed significantly from uniform ( $p < 0.001$  in all cases). To clarify the structure of the final distribution, we superimposed symbols (crosses) at rhythms whose intervals are related by simple integer ratios (Figure 2C). We plot all 22 simple ratios ( $i : j : k$ ) where  $1 \leq i, j, k \leq 3$ .

The modes of the distribution obtained in the last iteration closely correspond to integer ratio points, though not all integer ratio rhythms correspond to modes. We defined an “integer score” as the distance (JSD) between the data distribution and a uniform distribution over the 22 integer ratio points. To evaluate statistical significance, we compared the score to that of random distributions containing the same number of modes, revealing a significant tendency toward probability mass at integer ratios ( $p < 0.001$ ). To assess how closely the modes were centered on integer ratios, we fit a 22-component Gaussian mixture model to the data from the last iteration and measured whether the component means aligned with integer ratios. Two of the modes were significantly biased away from integer ratios (those corresponding to 2:3:1,  $p < 0.001$ ; those corresponding to 3:2:1,  $p = 0.002$ ), consistent with previous evidence for human rhythm categories centered close to, but not exactly on, integers [22].

We performed a similar experiment for two-interval rhythms, the advantage being that the lower dimensionality permitted



**Figure 1. Illustration of Stimuli and Experimental Paradigm**

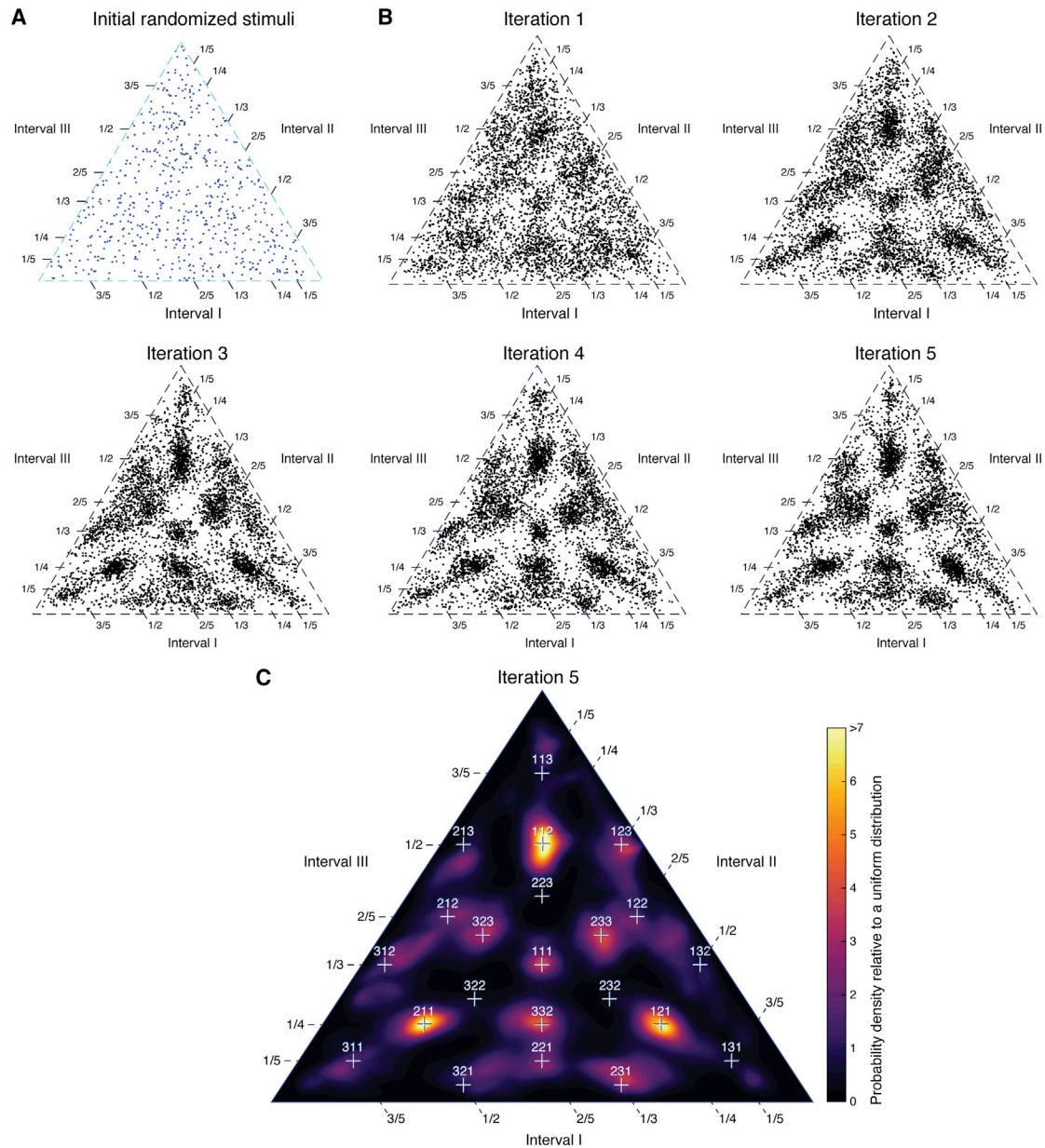
- (A) One repetition of a three-interval rhythm (four clicks) with integer ratio proportions (1:1:2), resulting in a rhythm with intervals of 25%, 25%, and 50% of the total pattern duration.  
 (B) One repetition of a three-interval rhythm with non-integer ratio proportions (1.6:2.14:1).  
 (C) Three repetitions of the 1:1:2 three-interval rhythm (ten clicks).  
 (D) The repeated pattern 1:1:2 written in Western musical notation.  
 (E) As first described in [20], three-interval rhythms can be presented as a point in a three-dimensional space, where the x, y, and z coordinates correspond to the first, second, and third interval, respectively. The points with an overall fixed duration (2,000 ms in this case) form a triangle (simplex) that can be projected to a two-dimensional plane, creating a “rhythm simplex,” also known as a “chronotopological map,” or “rhythm chart” [20, 27]. Here and in (F), the dashed triangle denotes the region from which stimuli were drawn (defined by inter-onset intervals that are larger than 15% of the total duration, e.g., 300 ms for the 2,000 ms duration used in most experiments), with two example integer ratio rhythms indicated by crosses.  
 (F) The rhythm simplex. The red line plots an example stimulus/response trajectory across the five iterations of an example trial.  
 (G) Schematic of trial structure.

ten iterations per seed, which we used to further examine convergence (Figure S1). As with three-interval rhythms, reproductions converged to a distribution whose modes were approximately simple integer ratios (1:1, 1:2, 1:3...), and copying accuracy converged by iteration 5 (see Figure S2 for convergence behavior for all experiments). It thus seems that the iterative process for simple rhythms converges after about five iterations (see also Figures S7C and S7D).

### Controlling for Productive Biases

Our method has the advantage of being efficient, comprehensive, and hypothesis neutral, but it relies on a complex task involving perception and synchronized reproduction, and the re-

sults could in principle reflect a combination of perceptual and productive constraints. To interpret the results in terms of a perceptual prior, we must ensure that production biases do not substantially contribute. We thus conducted a series of control experiments. To test whether the results depended on production modality, we asked participants to produce a vocal syllable rather than tap (experiment 2). To test whether the results depended on entrainment, we asked participants to replicate a pattern from memory rather than tap in synchrony (experiment 3). To test the role of successive iterations within participants, we temporally intermixed reproductions from different seeds and mixed iterative chains across participants (experiment 4). To verify that the predicted prior would be evident without



**Figure 2. Results of Experiment 1: Iterated Reproduction by Synchronous Tapping**

(A) All seeds of experiment 1 displayed on the rhythm simplex.

(B) All responses from each of the five iterations of experiment 1.

(C) Kernel density estimate of the continuous distribution underlying the data from iteration 5 of experiment 1. Crosses plot simple integer ratio rhythms.

See also [Figures S1](#), [S2](#), and [S4](#).

reproduction demands, we employed a standard discrimination task (experiment 5) and compared the results to a Bayesian ideal observer using the distribution obtained in experiment 1. Finally, to test whether integer ratios would result from reproducing any type of temporal pattern, we asked listeners to repeat the syllables of meaningful English phrases (experiment 6).

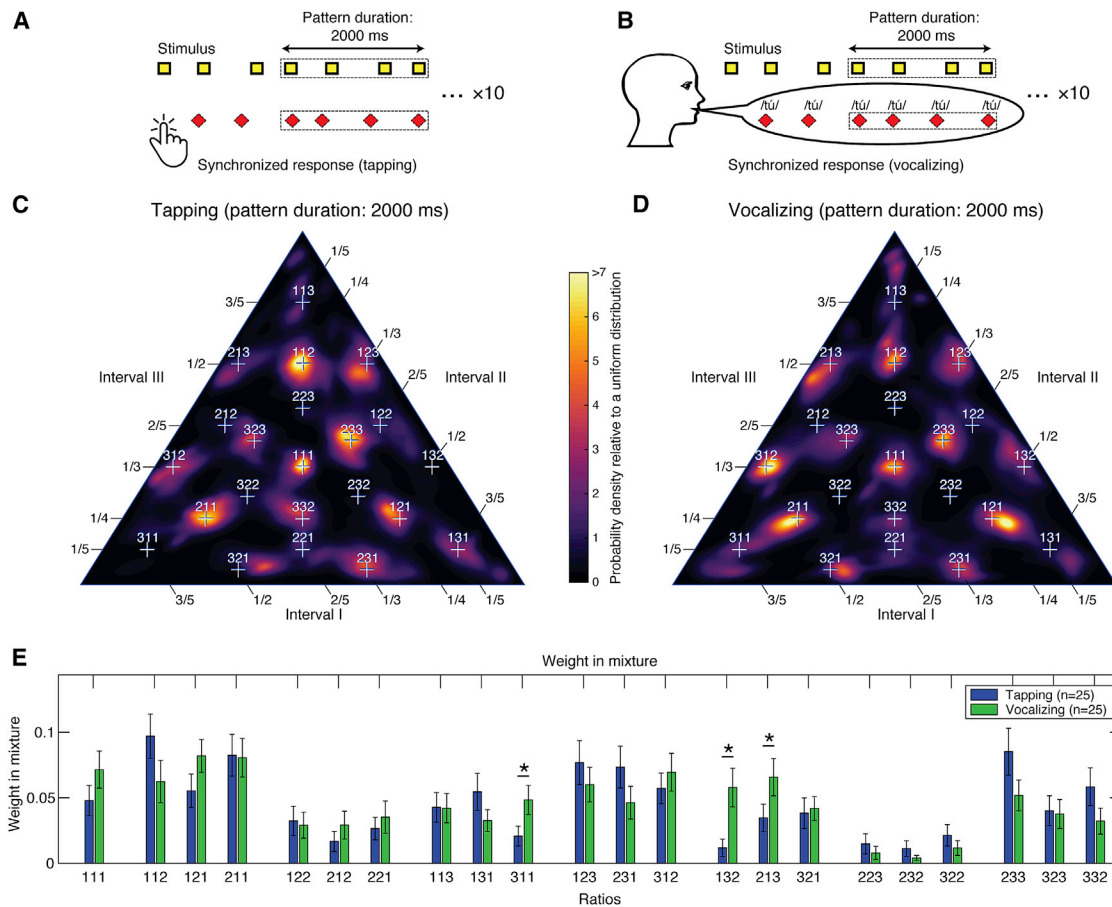
### Experiment 2: Tapping versus Vocalizing

In separate, counterbalanced blocks, participants reproduced sequences of clicks either by tapping as in experiment 1 ([Fig-](#)

[ure 3A](#)) or by vocalizing the syllable /tú/ ([Figure 3B](#)). The motivation was that production biases for speech and finger tapping might be expected to be different—speech rhythms are often produced aperiodically [29], whereas finger tapping is often periodic due to finger movement kinematics [14]. One might thus expect the results to differ with reproduction modality if production biases play a large role.

[Figures 3C](#) and [3D](#) show that tapping and vocalizing produce qualitatively similar distributions, with prominent integer ratio peaks in both cases (both integer scores were significant,





**Figure 3. Results of Experiment 2: Tapping versus Vocalizing**

(A) Schematic of the tapping task.  
 (B) Schematic of the vocalization task.  
 (C) Response distribution from the fifth iteration of the tapping task.  
 (D) Response distribution from the fifth iteration of the vocalization task.  
 (E) Weights of Gaussian mixture components assigned to each integer ratio rhythm in the two distributions (see also Figure S7). Error bars plot confidence intervals on the weights (SDs of the weight distributions derived from bootstrapping), with significance here and elsewhere evaluated after Bonferroni correction for multiple comparisons.

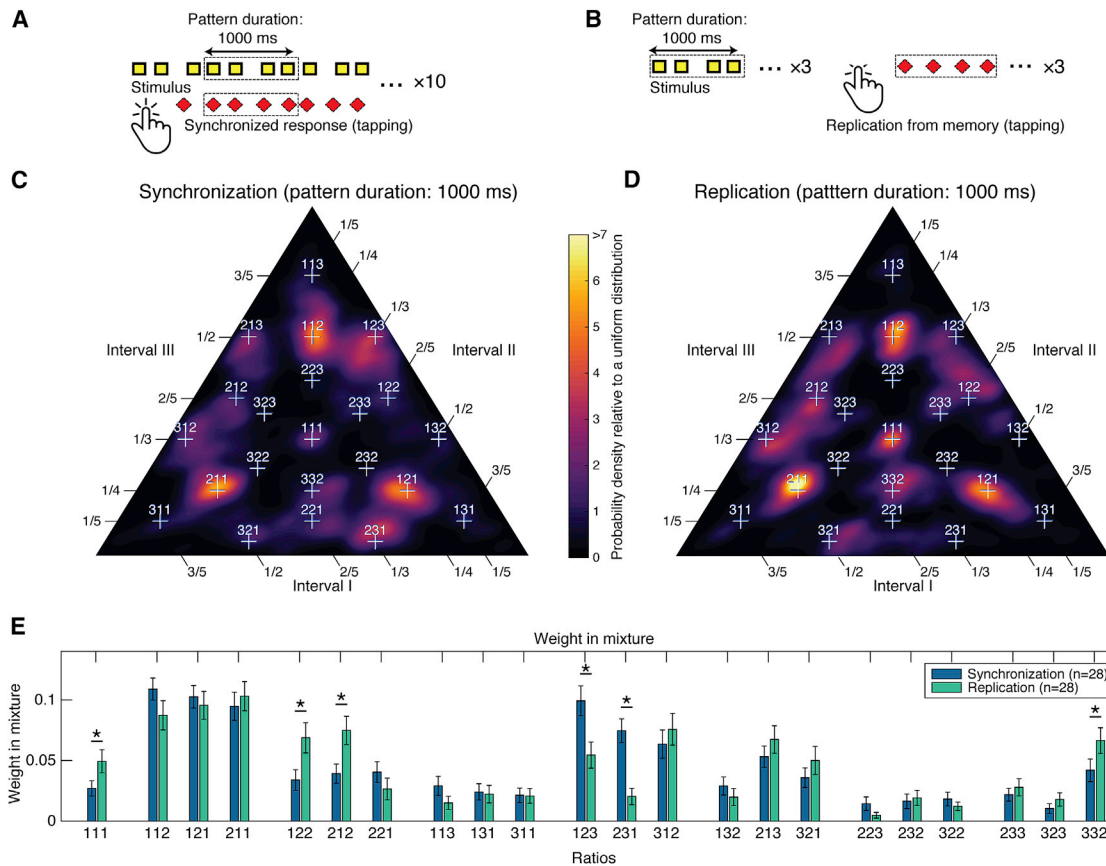
$p < 0.001$ ). To compare the detailed form of the distributions, we fit a Gaussian mixture model to the final (fifth) iteration for each condition, constraining the component means to the vicinity of a corresponding integer ratio point (Figures S7E–S7H demonstrate the resulting quality of fit). The weights of the mixture components for the two conditions were significantly correlated (Figure 3E;  $r = 0.58$ ,  $p = 0.005$ ,  $r = 0.83$  after correcting for test-retest reliability of the weights), and only three of the 22 integer ratio points showed a significant difference between conditions (3:1:1, 1:3:2, and 2:1:3;  $p = 0.01$ , 0.0002, and 0.04, respectively). These small differences were reflected in a significant JSD (JSD = 0.082,  $p = 0.001$ ). Overall, however, the results indicate that reproduction modality has little effect on the obtained distributions.

### Experiment 3: Replication from Memory versus Synchronization

To test whether the results were independent of the requirement to synchronize the reproduction with the heard signal,

we asked participants to listen to a rhythm and replicate it from memory (Figures 4A and 4B). To reduce working memory load, we reduced the rhythm duration to 1 s, doubling the tempo from that in experiments 1 and 2. Participants completed a separate block in which they tapped synchronously, as in experiment 1.

We observed a slightly different distribution compared to experiment 1, presumably due to the shorter duration (see Figure S3 for a separate experiment manipulating tempo), but both synchronization and replication showed significant integer scores ( $p < 0.001$ ). Moreover, integer ratio weights (Figure 4E) for the two conditions were correlated ( $r = 0.68$ ,  $p = 0.001$ ,  $r = 0.79$  after correcting for test-retest reliability of the weights). The differences between distributions could reflect the number of stimulus repetitions per iteration (ten for synchronization; three for replication, to reduce memory load). These results are consistent with findings of similar tapping biases for synchronization and replication from memory [21] and are consistent with the idea that our results reflect a perceptual prior.



**Figure 4. Results of Experiment 3: Synchronization and Replication**

(A) Schematic of the synchronization task.

(B) Schematic of the replication task.

(C) Response distribution from the fifth iteration of the synchronization task.

(D) Response distribution from the fifth iteration of the replication task.

(E) Weights of Gaussian mixture components assigned to each integer ratio rhythm in the two distributions. Error bars plot confidence intervals on the weights (SDs of the weight distributions derived from bootstrapping).

See also [Figure S3](#).

### Experiment 4: Eliminating Dependencies across Iterations

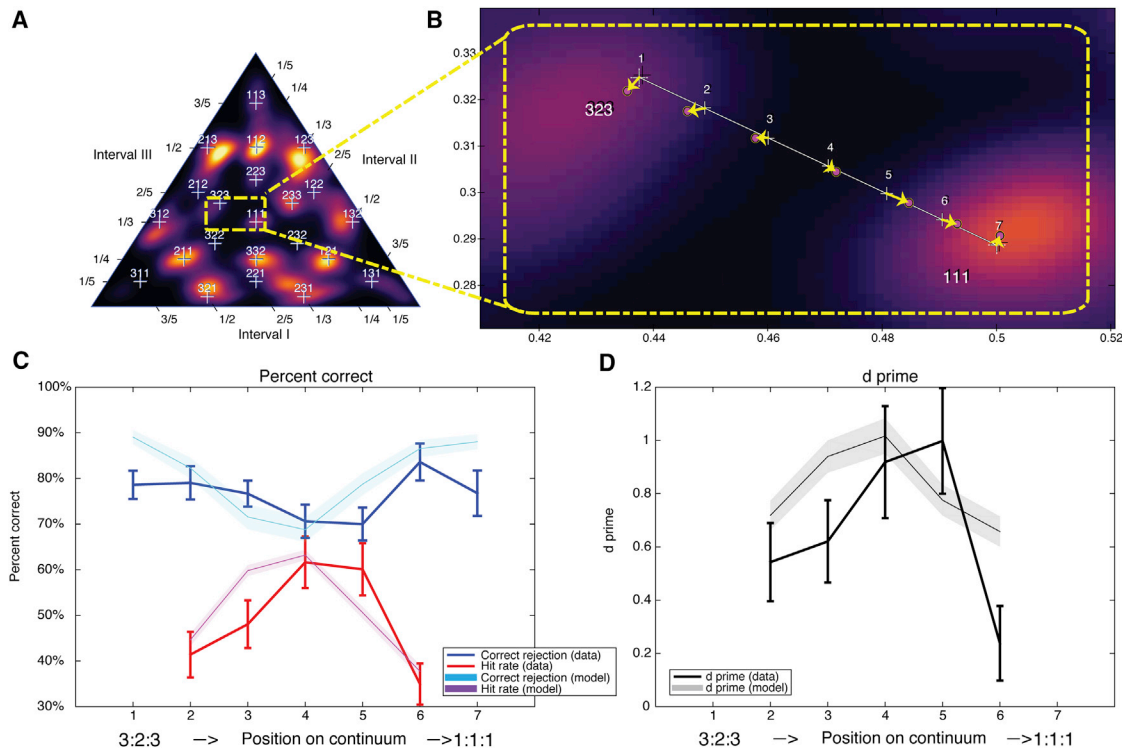
In order for iterated reproduction to converge to the prior, the reproductions in principle must be influenced only by the prior and the stimulus (the seed, or the previous iteration's reproduction), and not by previous history [30]. Our main paradigm could violate this assumption of Markovian behavior because the same participants complete all iterations for a particular seed in succession, and reproductions could be affected by the memory of previous iterations. To test the effect of such factors, we repeated experiment 1 with two alternative procedures. In the first condition ("random order within participants"), participants again completed five iterations per seed, but the iterations for different seeds were intermingled. In the second condition ("across participants"), each participant completed one iteration for a large number of seeds, and the reproductions of one participant formed the seed for the next participant ([Figure S4](#)).

The results of both conditions were similar and did not differ substantially from those of experiment 1 ([Figure S4](#)). Integer ratio weights were significantly correlated with those of experiment 1

in both conditions (random order within participant:  $r = 0.83$ ,  $p < 0.001$ ,  $r = 0.96$  after correcting for reliability; across participants:  $r = 0.76$ ,  $p < 0.001$ ,  $r = 0.85$  after correction), and in only two cases were individual weights significantly different. It thus appears that any effects of across-iteration memory are modest and that our basic paradigm is reasonable to employ in practice (the advantage being that it is logistically simpler to perform the experiment separately for each subject). These results also provide evidence that the priors being measured are similar across participants—if there was substantial variation, the two procedures would yield different results [31].

### Experiment 5: Rhythm Discrimination

To further test whether the results reflect perceptual priors, we used a discrimination task [19, 32] as is traditionally employed to probe priors. The disadvantage of this approach is that it is in practice not feasible to explore the entire space of rhythms. However, a classical discrimination task removes all production constraints and provides a useful consistency check. We measured discrimination of stimuli drawn from a



**Figure 5. Results of Experiment 5: Rhythm Discrimination, with Predictions from the Measured Prior**

(A) The response distribution from experiment S2 (with a pattern duration of 1,000 ms).

(B) Close-up view of a slice through rhythm space occupied by experimental stimuli. Stimuli are depicted with white crosses, and perceptual biases predicted by the prior are marked as magenta diamonds; yellow arrows show direction of bias. Stimuli are “pulled” toward the modes, such that estimates of points straddling the trough are further apart than those of points on one side of the trough, improving discrimination.

(C) Hit and correct rejection rates for human listeners and for an ideal observer using the measured prior. Error bars plot the SEM.

(D) Perceptual sensitivity for each position along the continuum, along with the predicted sensitivity of the ideal observer from (C). The value plotted for position  $k$  is the discriminability of the rhythms at positions  $k - 1$  and  $k + 1$ . Values are thus available only for positions 2 through 6. Error bars plot the SEM.

See also [Figure S7](#).

one-dimensional slice through the three-interval rhythm simplex. If our results indeed reflect a prior, discrimination should be best at the troughs of the distribution and worst at the modes, as in classic categorical perception (because under Bayesian inference, estimates of stimuli are typically “pulled” toward the closest mode of the prior, rendering pairs of stimuli close to a mode more similar than pairs of stimuli between modes [11]).

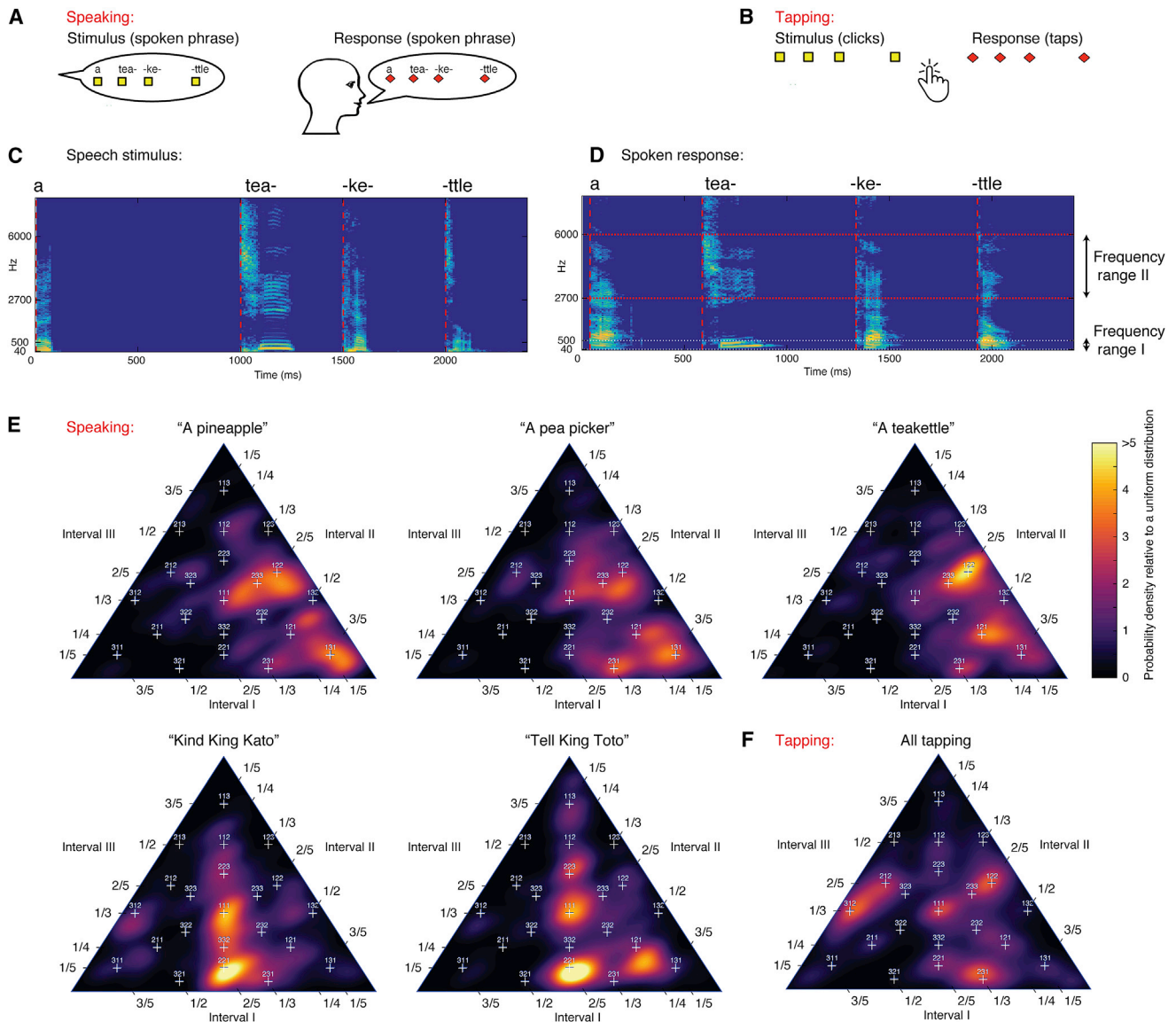
Participants discriminated pairs of three-interval rhythms straddling positions along a continuum between 1:1:1 and 3:2:3 (Figures 5A and 5B). Consistent with previous discrimination results [19, 32], and as predicted by a Bayesian interpretation of our iterated reproduction results, discriminability increased between the modes of the distribution measured in experiment 1 (Figures 5C and 5D;  $F(4,18) = 10.77$ ,  $p < 0.001$ ). A planned comparison between the average discriminability of stimuli closest to integer ratio rhythms (points 2 and 6 in Figure 5D) and of stimuli in between (points 3–5) showed better discrimination for the former, as predicted ( $t(18) = 6.27$ ,  $p < 0.001$ ). Moreover, the discrimination performance of a Bayesian ideal observer using the prior obtained in experiment 1 qualitatively replicated human performance (Figures 5C and 5D). These results provide further evidence that our method produces the perceptual prior for rhythm.

### Experiment 6: Tapping and Speaking

To test whether integer ratio biases extend to the perception of all sound signals, we asked participants to repeat spoken utterances in an analogous iterated reproduction paradigm (see also [33]). The syllables of five four-syllable phrases were segmented from recordings and substituted for the clicks used in our main paradigm (i.e., presented with initially random time intervals between syllables); participants were asked to repeat them back (Figure 6A). The onsets of the replicated syllables were detected and substituted for the initial onsets times as in experiments 1–4 (Figures 6C and 6D). We chose phrases with distinctive rhythms and short syllables with salient onsets. To facilitate syllable pronunciation without overlap or coarticulation, we presented phrases at a relatively slow tempo. Participants also completed the analogous task with clicks and taps (Figure 6B) in a separate block.

In contrast to click rhythms, speech rhythms converged to distributions whose peaks did not overlap with integer points. None of the five spoken patterns showed a significant integer score ( $p = 0.07, 0.1, 0.1, 0.85$ , and  $0.14$  for phrases 1–5, respectively). In contrast, the same participants produced significant integer scores when asked to replicate click rhythms by tapping (Figure 6D;  $p < 0.001$  for all five equal-sized subsets of tapping data, analyzed in subsets to equate in power to individual speech





**Figure 6. Results of Experiment 6: Speaking versus Tapping**

(A) Schematic of the speaking task. Listeners heard a spoken phrase, segmented into syllables that were assigned random onset times, and repeated it.

(B) Schematic of the tapping task.

(C) Spectrogram of an example speaking stimulus. Silences of variable durations were inserted between syllables such that the intervals between voicing onsets (dashed red lines) had the desired intervals—in this example,  $r_1 = 1000$ ,  $r_2 = 500$ , and  $r_3 = 500$  ms (for illustrative purposes).

(D) Spectrogram of example recorded response. Dotted horizontal lines mark the two ranges used for the onset detection process (range I, white: 40–500 Hz; range II, red: 2700–6000 Hz). Vertical dashed lines mark the detected onsets.

(E) Response distribution from the fifth iteration of the speaking task, for each of five phrases.

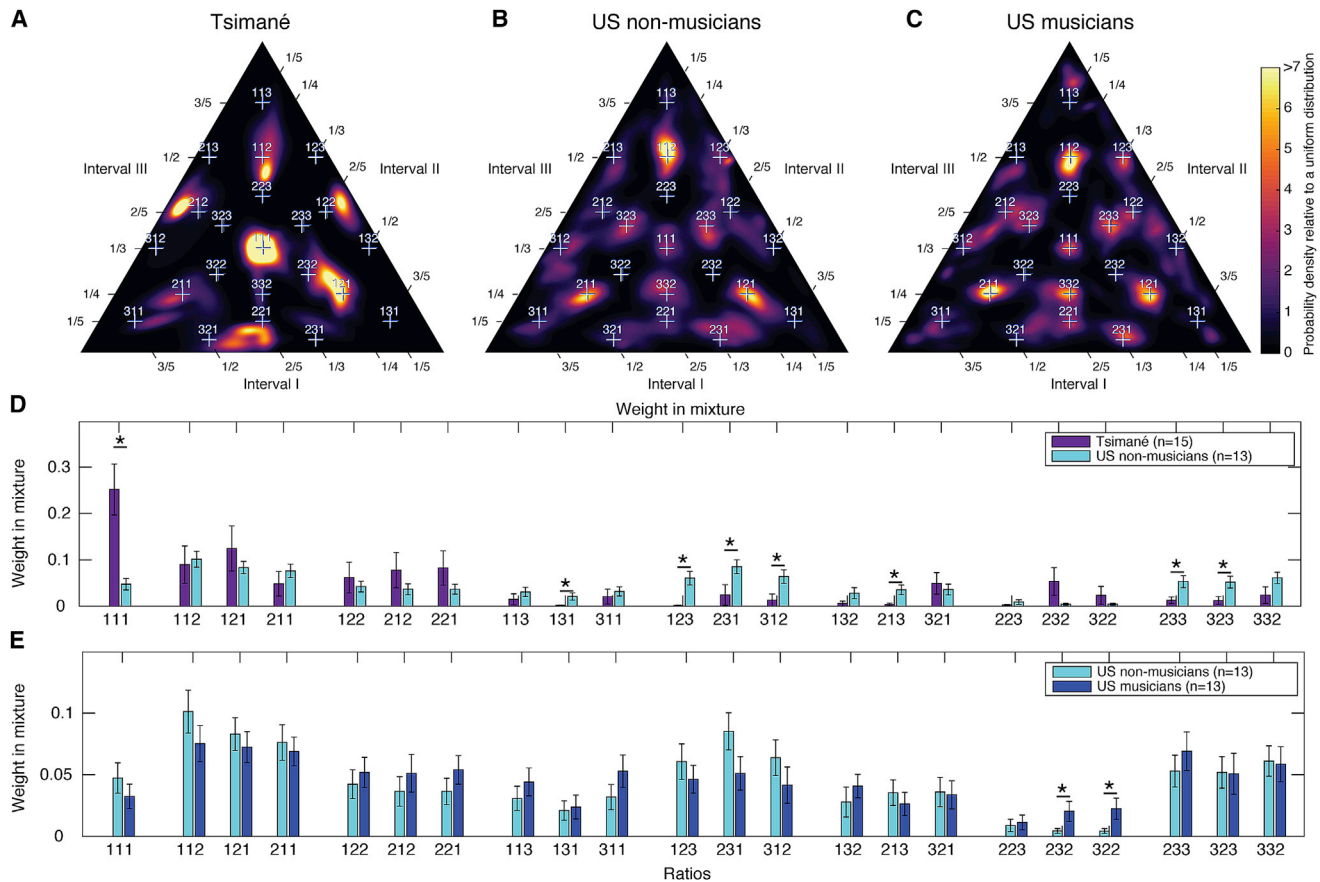
(F) Response distribution from the fifth iteration of the tapping task.

See also [Figures S5](#) and [S6](#).

conditions). These spoken phrases were a lone exception among our experimental conditions; all others produced statistically significant integer scores ([Figure S6](#)). The results indicate that the primacy of integer ratio rhythms is at least somewhat specific to music (consistent with well-known differences between speech and music rhythm [29, 34]) and underscore that integer ratios do not always emerge from iterated reproduction.

The results for both tasks are asymmetric; this is most likely because the stimulus was only presented once per iteration (all

other experiments used cyclically repeated stimuli). This effect is consistent with previous findings that tapping biases are altered by cyclic repetition [21]. Because cyclic permutations of patterns (e.g., 1:1:2 and 1:2:1) produce stimulus sequences that are identical but for the beginning and end, repetition could cause these patterns to be similarly perceived, reducing the differences in the measured prior for distinct cyclic permutations. However, we found that adding an accent to the first click of the repeating stimulus (to remove this ambiguity) had little effect



**Figure 7. Priors in Amazonians and US Participants with and without Musical Training**

(A–C) Response distribution from fifth iteration of tapping task for Tsimané participants (A), US non-musicians (B), and US musicians (C). (D and E) Weights of Gaussian mixture components assigned to each integer ratio rhythm for Tsimané and age-matched US participants with little musical exposure (D) and US musicians and non-musicians (E). Error bars plot confidence intervals on the weights (SDs of the weight distributions derived from bootstrapping).

on the results (Figure S5), suggesting that cyclic ambiguity is not the main reason for the symmetry of many repetitions compared to one. Instead, it seems that repetition itself may change the prior probability of rhythmic patterns, consistent with its importance in music [35].

### Experiment 7: Priors in Native Amazonians

To explore the possible dependence of rhythmic biases on exposure to music, we tested members of the Tsimané, an indigenous society of farmer-foragers residing in remote areas of the Bolivian Amazon basin. The Tsimané remain relatively isolated from Western music and exhibit dramatically different harmony preferences than US participants [36], raising the possibility that they might exhibit distinct rhythmic priors as well. None of the Tsimané participants regularly played a musical instrument. We therefore compared them to an age-matched group of US non-musicians. Despite their profoundly different cultural exposure, the Tsimané also displayed a significant integer score ( $p = 0.005$ ). However, the prior derived from their reproductions differed substantially from that of US participants (Figures 7A and 7B; JSD = 0.23,  $p < 0.001$ ). Although both cultures assigned high probability to some integer ratio rhythms (e.g., 1:1:2 and

1:2:2), eight of the 22 rhythms showed significant differences between cultures, and the integer ratio weights of the two populations were not significantly correlated ( $r = 0.19$ ,  $p = 0.43$ ; Figure 7D).

Several uninteresting explanations for the cross-cultural differences are unlikely given the constraints of our method and the results. It is not the case that the Tsimané simply cannot entrain to rhythms, because the method discards taps that do not closely correspond to stimulus events. It is also not the case that the differences reflect an inability of the Tsimané to reproduce short intervals, because three of the modes of their distribution lie near the boundary of the rhythm simplex, and as such they contain intervals close to the shortest allowable in our paradigm. And although Tsimané tapping was more variable than that of US participants, evident in the broader modes of their response distribution, it is visually apparent that there is “room” for additional modes on the rhythm simplex that nonetheless did not arise.

The Tsimané prior is consistent with previous ethnographic descriptions of their music ([37], pp. 345–368), which reported patterns of 1:1:1 and 1:1:2 (and their cyclic permutations). In contrast, patterns prominent in US participants but not in

Tsimané, e.g., 2:3:3 and 1:2:3, are not evident in descriptions of Tsimané music. Like that of the US participants, the Tsimané distribution showed significant deviations from integer ratios in some cases (1:2:2,  $p = 0.005$ ; 2:1:2,  $p < 0.001$ ).

### Effect of Musical Expertise

The apparent effects of experience found in experiment 7 raise the question of whether priors primarily reflect passive experience with music. To explore this issue, we compared US musicians and non-musicians (Figures 7B and 7C). Although the two resulting distributions were significantly different (JSD = 0.08,  $p < 0.001$ ), it is visually apparent that the distributions are qualitatively similar. Both groups exhibited a significant integer score ( $p < 0.001$ ), and their integer ratio weights were correlated ( $r = 0.79$ ,  $p < 0.001$ ,  $r = 0.91$  after correcting for reliability; Figure 7E). Out of the 22 simple integer ratio rhythms, only two were more prominent in musicians than non-musicians (2:3:2 and 3:2:2). These patterns occur in Balkan music but are not accurately perceived by Western adults and children [38]. The difference in weights for these rhythms could be greater familiarity of musicians with music patterns from other cultures.

It is also apparent that the effect of culture was more prominent than the effect of musical training; the JSD between Tsimané and US non-musicians (JSD = 0.23) was larger than the JSD between US musicians and non-musicians (JSD = 0.08,  $p = 0.02$ ). Overall, these results indicate that rhythm priors are not strongly affected by musical expertise, in US participants at least, but that culture has a significant role in shaping musical priors.

## DISCUSSION

We have introduced a method to measure internal biases in rhythm reproduction and have presented evidence that the method reveals perceptual priors on rhythm that constrain what listeners can accurately perceive and reproduce. Our method permits the measurement of the prior over the entire space of simple rhythms, and as such it could reveal priors no matter what form they take. We found that priors on rhythm in US participants favored integer ratios, but only those that are prevalent in Western music. Priors in a native Amazonian society also exhibited peaks at integer ratios, but they were otherwise qualitatively different from priors in US participants, in ways that are consistent with what we know of their music. The results were similar for several different modes of reproduction but did not extend to the reproduction of spoken phrases, indicating that integer ratio priors are at least somewhat specific to music. In addition, the results were similar for musicians and non-musicians, suggesting that priors are shaped by passive exposure to music.

### Relation to Other Iterative Methods

Our general methodology is a version of the serial reproduction paradigm first introduced by Bartlett [23] to demonstrate cultural biases in the perception of short stories, jokes, and pictures. Serial reproduction was subsequently used as a model for the evolution of language through cultural transmission (“iterated learning”; [10, 26]). Interest in iterated paradigms has grown recently due to their mathematical similarity to sampling tech-

niques in modern machine learning [10]. Serial reproduction has since been used to study visual biases on color naming [6], the emergence of linguistic structure during transmission of language [26], biases in speech intonation and whistling [25, 33], and emergent structure in symbolic sequences [31] and non-verbal gestural cues [39]. These previous approaches either have used discrete language-mediated judgments or did not systematically map a perceptual space. To our knowledge, our method is the first application of serial reproduction to map an entire analog perceptual space.

### Previous Evidence for Rhythmic Priors

Although previous research on musical rhythm has generally not made reference to a prior or to Bayesian inference (but see [18]), several previous findings are consistent with our results. Discrimination of simple rhythmic stimuli has been found to be better near boundaries between simple integer rhythms [19, 32], a finding interpreted as evidence for categorical perception. Our discrimination results are qualitatively similar to (and were inspired by) these prior findings, but they are predicted by the prior measured using iterative reproduction, suggesting that categorical perception can result from rhythm priors. Our contribution relative to this discrimination literature lies in providing a method that is more efficient in mapping perceptual biases (enabling high-resolution estimates of the prior). We quantified the efficiency benefit by estimating the number of discrimination trials required to estimate the three-interval prior with the fidelity obtained by our iterative method. We estimate that our iterative procedure recovers the prior 20–40 times faster than traditional discrimination experiments (see the [Supplemental Experimental Procedures](#)). Obtaining results equivalent to those of experiment 1 using discrimination would require more than 520 participants (assuming the same session duration).

Another approach to understanding mental representations of rhythm has been to ask trained musicians to categorize rhythms using musical notation. Desain and Honing [20] applied this approach to three-interval rhythms; the results were later used to infer the prior probability of particular integer ratios [18]. Our results are consistent with their findings—rhythms associated with musically notated categories were close to the modes found by our implicit paradigm. The benefit of our method is that it is independent of Western musical notation and can thus be used to test non-musicians or non-Western listeners.

Our results can be viewed as an extension of tapping reproduction experiments, in which reproductions of two- or three-interval rhythms are biased toward integer ratios [16, 21, 22]. However, because the production biases measured in a single reproduction are small, these experiments occasionally provided inconsistent results [22] and did not provide a way to reveal the prior. The first iteration of our basic paradigm is equivalent to such an experiment, but the reproduction biases are enhanced with iteration (Figures 2 and S2), enabling direct measurement of the prior. Our experiments also reveal that the underlying effects are largely independent of production constraints—tapping is merely a way to access a listener’s representation of rhythm.

### Factors Influencing Rhythmic Priors

Although we have argued that our method reveals a listener’s prior over rhythms, the detailed form of the obtained distribution

varied across several experiments. Task demands required across-experiment differences in tempo, as well as the number of stimulus repetitions per iteration. It seems plausible that both factors affect the prior probability of rhythms. Some patterns could be more common in particular tempo ranges, consistent with evidence that rhythm perception is modulated by tempo [28, 40]. And repetition could affect the tendency for patterns to be interpreted as musical [35, 41], which might amplify the bias for integer ratio rhythms.

### Origins and Functions of Priors on Rhythm

One reason to be interested in priors on rhythm (apart from the role they likely play in extracting musical structure from sound, particularly from multi-instrument pieces [12]) is that they constrain what can be accurately perceived and transmitted. As such, if priors were themselves tightly constrained, these constraints could in principle explain universal features of rhythm. Previous proposals for such biological and/or cognitive constraints have included oscillatory circuits [15, 42, 43], recurrent attractor networks [44], and rhythm representations built on an isochronous pulse [16], all of which should bias perception toward integer ratio rhythms. Alternatively (or additionally), the widespread presence of integer ratio rhythms could reflect constraints on production [14, 45].

Our results do not resolve these issues, but the differences between cultures that we observed indicate that rhythm priors are not fully fixed by biological constraints. The results leave open the possibility that priors themselves are largely free to reflect whatever distribution occurs in the world [5], with cross-cultural similarities instead being driven by productive or other constraints on human musical systems. One informative avenue will be to document the prior over development, perhaps by applying our methods to children.

Further insights into these issues could also be gained by studying additional cultures, particularly those whose musical rhythms deviate substantially from integer ratio proportions [46–48]. Our paradigm is ideal for this purpose. Corpus analysis is no substitute: notated music does not exist for many non-Western cultures, and when available, notations usually reflect what a Westerner hears when listening to the other culture's music. Moreover, it remains possible that priors could deviate from the empirical distribution of a culture's music. Our paradigm instead directly probes the mental representation of the listener. The results here provide another example of the perceptual diversity that can be revealed by experiments in non-Western cultures [36, 49, 50].

### EXPERIMENTAL PROCEDURES

See the [Supplemental Experimental Procedures](#) for a full description.

### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and seven figures and can be found with this article online at <http://dx.doi.org/10.1016/j.cub.2016.12.031>.

### AUTHOR CONTRIBUTIONS

N.J and J.H.M designed the experiments and wrote the paper. N.J collected the data.

### ACKNOWLEDGMENTS

The authors thank R. Godoy, T. Huanca, E. Gibson, and J. Canfield for logistical assistance; M. Salinas, V. Zhao, K. Woods, and R. Futrell for assistance with experiments; T. Griffiths for advice and support; and R. Polak, B. H. Repp, J. London, C. Raz, H. Honing, and the McDermott Lab for comments on the manuscript. This work was supported by a Hebrew University ELSC Postdoctoral Fellowship to N.J., the Center for Society and Neuroscience at Columbia University, and a McDonnell Scholar Award to J.H.M. All participants provided informed consent in accordance with the Massachusetts Institute of Technology Committee on the Use of Humans as Experimental Subjects.

Received: August 18, 2016

Revised: November 18, 2016

Accepted: December 13, 2016

Published: January 5, 2017

### REFERENCES

1. Knill, D.C., and Richards, W., eds. (1996). *Perception as Bayesian Inference* (Cambridge University Press).
2. Adams, W.J., Graf, E.W., and Ernst, M.O. (2004). Experience can change the 'light-from-above' prior. *Nat. Neurosci.* 7, 1057–1058.
3. Weiss, Y., Simoncelli, E.P., and Adelson, E.H. (2002). Motion illusions as optimal percepts. *Nat. Neurosci.* 5, 598–604.
4. Stocker, A.A., and Simoncelli, E.P. (2006). Noise characteristics and prior expectations in human visual speed perception. *Nat. Neurosci.* 9, 578–585.
5. Girshick, A.R., Landy, M.S., and Simoncelli, E.P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nat. Neurosci.* 14, 926–932.
6. Xu, J., Dowman, M., and Griffiths, T.L. (2013). Cultural transmission results in convergence towards colour term universals. *Proc. Biol. Sci.* 280, 20123073.
7. Jazayeri, M., and Shadlen, M.N. (2010). Temporal context calibrates interval timing. *Nat. Neurosci.* 13, 1020–1026.
8. Fischer, B.J., and Peña, J.L. (2011). Owl's behavior and neural representation predicted by Bayesian inference. *Nat. Neurosci.* 14, 1061–1066.
9. Körding, K.P., and Wolpert, D.M. (2006). Bayesian decision theory in sensorimotor control. *Trends Cogn. Sci.* 10, 319–326.
10. Griffiths, T.L., and Kalish, M.L. (2007). Language evolution by iterated learning with bayesian agents. *Cogn. Sci.* 31, 441–480.
11. Feldman, N.H., Griffiths, T.L., and Morgan, J.L. (2009). The influence of categories on perception: explaining the perceptual magnet effect as optimal statistical inference. *Psychol. Rev.* 116, 752–782.
12. McDermott, J.H. (2009). The cocktail party problem. *Curr. Biol.* 19, R1024–R1027.
13. Savage, P.E., Brown, S., Sakai, E., and Currie, T.E. (2015). Statistical universals reveal the structures and functions of human music. *Proc. Natl. Acad. Sci. USA* 112, 8987–8992.
14. Maes, P.J., Leman, M., Palmer, C., and Wanderley, M.M. (2014). Action-based effects on music perception. *Front. Psychol.* 4, 1008.
15. Large, E.W., and Snyder, J.S. (2009). Pulse and meter as neural resonance. *Ann. N Y Acad. Sci.* 1169, 46–57.
16. Povel, D.J. (1981). Internal representation of simple temporal patterns. *J. Exp. Psychol. Hum. Percept. Perform.* 7, 3–18.
17. Temperley, D. (2007). *Music and Probability* (The MIT Press).
18. Sadakata, M., Desain, P., and Honing, H.J. (2006). The Bayesian way to relate rhythm perception and production. *Music Percept.* 23, 269–288.
19. Clarke, E.F. (1987). Categorical rhythm perception: an ecological perspective. In *Action and Perception in Rhythm and Music, Volume 55*, A. Gabrielsson, ed. (Royal Swedish Academy of Music), pp. 19–33.
20. Desain, P., and Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception* 32, 341–365.



21. Repp, B.H., London, J., and Keller, P.E. (2013). Systematic distortions in musicians' reproduction of cyclic three-interval rhythms. *Music Percept.* *30*, 291–305.
22. Repp, B.H., London, J., and Keller, P.E. (2012). Distortions in reproduction of two-interval rhythms: when the “attractor ratio” is not exactly 1:2. *Music Percept.* *30*, 205–223.
23. Bartlett, F.C. (1932). *Remembering: A Study in Experimental and Social Psychology* (Cambridge University Press).
24. Xu, J., and Griffiths, T.L. (2010). A rational analysis of the effects of memory biases on serial reproduction. *Cognit. Psychol.* *60*, 107–126.
25. Verhoef, T., Kirby, S., and de Boer, B. (2014). Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *J. Phonetics* *43*, 57–68.
26. Kirby, S., Cornish, H., and Smith, K. (2008). Cumulative cultural evolution in the laboratory: an experimental approach to the origins of structure in human language. *Proc. Natl. Acad. Sci. USA* *105*, 10681–10686.
27. Honing, H. (2013). The structure and interpretation of rhythm in music. In *The Psychology of Music, Third Edition*, D. Deutsch, ed. (Academic Press), pp. 369–404.
28. London, J. (2012). *Hearing in Time: Psychological Aspects of Musical Meter* (Oxford University Press).
29. Turk, A., and Shattuck-Hufnagel, S. (2014). Timing in talking: what is it used for, and how is it controlled? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* *369*, 20130395.
30. Griffiths, T.L., Christian, B.R., and Kalish, M.L. (2008). Using category structures to test iterated learning as a method for identifying inductive biases. *Cogn. Sci.* *32*, 68–107.
31. Claidière, N., Smith, K., Kirby, S., and Fagot, J. (2014). Cultural evolution of systematically structured behaviour in a non-human primate. *Proc. Biol. Sci.* *281*, 20141541.
32. James, S., and Marvin, E. (2015). Categorical perception of irregular rhythms. In *2015 Annual Conference of the Society for Music Perception and Cognition* (Vanderbilt).
33. Braun, B., Kochanski, G., Grabe, E., and Rosner, B.S. (2006). Evidence for attractors in English intonation. *J. Acoust. Soc. Am.* *119*, 4006–4015.
34. Patel, A.D. (2010). *Music, Language, and the Brain* (Oxford University Press).
35. Margulis, E.H. (2014). *On Repeat: How Music Plays the Mind* (Oxford University Press).
36. McDermott, J.H., Schultz, A.F., Undurraga, E.A., and Godoy, R.A. (2016). Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature* *535*, 547–550.
37. Riestler, J. (1978). *Canción y Producción en la Vida de un Pueblo Indígena: Los Chimane del Oriente Boliviano* (Editorial Los Amigos del Libro).
38. Hannon, E.E., and Trehub, S.E. (2005). Metrical categories in infancy and adulthood. *Psychol. Sci.* *16*, 48–55.
39. Perlman, M., Dale, R., and Lupyan, G. (2015). Iconicity can ground the creation of vocal symbols. *R. Soc. Open Sci.* *2*, 150152.
40. Repp, B.H., London, J., and Keller, P.E. (2005). Production and synchronization of uneven rhythms at fast tempi. *Music Percept.* *23*, 61–78.
41. Tierney, A., Dick, F., Deutsch, D., and Sereno, M. (2013). Speech versus song: multiple pitch-sensitive areas revealed by a naturally occurring musical illusion. *Cereb. Cortex* *23*, 249–254.
42. Bååth, R., Lagerstedt, E., and Gärdenfors, P. (2013). An oscillator model of categorical rhythm perception. In *Proceedings of the 35<sup>th</sup> Annual Conference of the Cognitive Science Society* (Cognitive Science Society), pp. 1803–1808.
43. Peper, C., Beek, P.J., and van Wieringen, P.C. (1995). Multifrequency coordination in bimanual tapping: Asymmetrical coupling and signs of supercriticality. *J. Exp. Psychol. Hum. Percept. Perform.* *21*, 1117.
44. Desain, P., and Honing, H., eds. (1989). *Quantization of Musical Time: A Connectionist Approach, Volume 1* (MIT Press).
45. Merchant, H., Zarco, W., Pérez, O., Prado, L., and Bartolo, R. (2011). Measuring time with different neural chronometers during a synchronization-continuation task. *Proc. Natl. Acad. Sci. USA* *108*, 19784–19789.
46. Polak, R. (2010). Rhythmic feel as meter: non-isochronous beat subdivision in jembe music from Mali. *Music Theory Online* *16*. <http://mto.societymusictheory.org/issues/mto.10.16.4/mto.10.16.4.polak.html>.
47. Friberg, A., and Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: evidence for a common rhythmic pattern. *Music Percept.* *19*, 333–349.
48. Haugen, M.R. (2014). Studying rhythmical structures in Norwegian folk music and dance using motion capture technology: a case study of Norwegian telespringar. *Musikk og Tradisjon* *28*, 27–52.
49. Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A.D., and Koelsch, S. (2009). Universal recognition of three basic emotions in music. *Curr. Biol.* *19*, 573–576.
50. Cameron, D.J., Bentley, J., and Grahn, J.A. (2015). Cross-cultural influences on rhythm processing: reproduction, discrimination, and beat tapping. *Front. Psychol.* *6*, 366.