Why Do Durations in Musical Rhythms Conform to Small Integer Ratios?

Andrea Ravignani [1,2,3]*, Bill Thompson [1,2], Massimo Lumaca [4†] and Manon Grube [4†]

[1] *Language and Cognition Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands,* [2] *Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium,* [3] *Research Department, Sealcentre Pieterburen, Pieterburen, Netherlands,* [4] *Department of Clinical Medicine, Center for Music in the Brain, Aarhus University, Aarhus, Denmark*

One curious aspect of human timing is the organization of rhythmic patterns in small integer ratios. Behavioral and neural research has shown that adjacent time intervals in rhythms tend to be perceived and reproduced as approximate fractions of small numbers (e.g., 3/2). Recent work on iterated learning and reproduction further supports this: given a randomly timed drum pattern to reproduce, participants subconsciously transform it toward small integer ratios. The mechanisms accounting for this "attractor" phenomenon are little understood, but might be explained by combining two theoretical frameworks from psychophysics. The scalar expectancy theory describes time interval perception and reproduction in terms of Weber's law: just detectable durational differences equal a constant fraction of the reference duration. The notion of categorical perception emphasizes the tendency to perceive time intervals in categories, i.e., "short" vs. "long." In this piece, we put forward the hypothesis that the integer-ratio bias in rhythm perception and production might arise from the interaction of the scalar property of timing with the categorical perception of time intervals, and that neurally it can plausibly be related to oscillatory activity. We support our integrative approach with mathematical derivations to formalize assumptions and provide testable predictions. We present equations to calculate durational ratios by: (i) parameterizing the relationship between durational categories, (ii) assuming a scalar timing constant, and (iii) specifying one (of K) category of ratios. Our derivations provide the basis for future computational, behavioral, and neurophysiological work to test our model.

Keywords: rhythm, music perception, scalar expectancy theory, neural oscillations, integer ratio

INTEGER RATIOS AND MUSICAL RHYTHM

What are *small integer ratios*, and what makes integer-ratio rhythms special? A *ratio* between two inter-onset-intervals (IOIs) is the division between two, usually adjacent durations. *Integer* ratios can be written as a fraction: 1.5 equals 15/10 or 3/2, but $\sqrt{2}$ for instance cannot be written as a fraction. An integer ratio is *small* if the result of the division can be written as a small integer number divided by another small integer number e.g., 2/3, but not 23/51 (Pikovsky et al., 2003; Strogatz, 2003).

A *rhythm*, by definition as used here, is a pattern of durations (London, 2004, p. 4) characterized by the succession of event onsets over time, in other words a series of IOIs.

Auditory rhythms with small integer ratios between IOIs are common in the world's music (Essens and Povel, 1985; Toussaint, 2013; Savage et al., 2015). Psychological and neural research suggests that small integer-ratio rhythms allow a more accurate internal representation (Essens, 1986; Sakai et al., 1999), improved deviance detection (Jones and Yee, 1997; Large and Jones, 1999), enhanced memory (Deutsch, 1986; Palmer and Krumhansl, 1990) and reproduction (Povel and Essens, 1985; Essens, 1986), and better synchronization (Patel et al., 2005). The distortion of near-integer ratios toward integer ones (or their harmonics) reported in behavioral (Fraisse, 1982) and neurophysiological studies (Motz et al., 2013) further supports the idea of small ratios acting as "attractors" (Gupta and Chen, 2016). This idea has recently received support from studies of iterated learning and reproduction. When humans reproduce an initially randomly-timed rhythmic sequence, and this process is repeated in a cascade fashion within one or across several individuals, the sequence is subconsciously reshaped to be composed of IOIs related by small integer ratios (**Figure 1A**; c.f. Polak et al., 2016; Ravignani et al., 2016, 2018; Jacoby and McDermott, 2017).

Why do rhythms (i.e., patterns of durations) tend to exhibit small integer ratios? Why are humans drawn to rhythms with such a peculiar mathematical property, in both perception and production? Does this property reflect a special quirk of music perception and/or motor sequencing, or could it be explained by domain-general aspects of cognition? Can we explore these alternatives through mathematical formalism? Here, we explore mathematically the possibility that the human bias toward small integer ratios may be explained by a combination of scalar expectancy and categorical perception.

We begin by outlining the relevant classical frameworks for human timing, and go on to summarize the evidence in support of the small-integer ratio bias in rhythm perception. We then present our proposal linking the frameworks to the bias through mathematical formalisms. Specifically, we draw on the scalar property of time interval estimation to formulate a simple model of categorical perception that may result in an integer ratio bias (**Figure 1**), and link this to neural oscillations. We conclude by briefly discussing the merits and limitations of our model and outlining future goals.

PSYCHOPHYSICAL AND OSCILLATORY APPROACHES

Two major theoretical approaches, among several, have been suggested to account for the mechanisms behind human timing (Wing and Kristofferson, 1973a,b; Getty, 1975; Meck, 1996; Church, 1999; Grondin, 2001, 2010; Mauk and Buonomano, 2004; Karmarkar and Buonomano, 2007; Ivry and Schlerf, 2008; Allman et al., 2014; Merker, 2014). The most influential and empirically tested psychoacoustic model is the "scalar expectancy theory" (Wearden, 1991; Allman and Meck, 2011). Psychophysical research shows that human timing often follows Weber's law (Bizo et al., 2006): the error for an interval duration being timed is proportional to the duration of that interval. One perception-based formulation states that the ratio between

the just-noticeable difference (JND) and the duration of a reference stimulus is constant across stimulus length (Grondin, 2001). In another formulation, the coefficient of variation (standard deviation divided by mean) in estimating durations is constant across durations (**Figure 1D**; Gibbon, 1977).

Another relevant approach to timing mechanisms comes from neuroscience and physics. It suggests that neural oscillations entrain (or even "resonate") with the periodicity of external stimuli at multiple time-scales (Buzsaki, 2006; Large, 2008; Arnal and Giraud, 2012; Gupta, 2014; Aubanel et al., 2016; Celma-Miralles et al., 2016). Specifically, it states that phase and frequency of neural oscillations entrain with the phase and frequency of external events at multiple metrical levels. For instance, processing a metronome beat will induce low-frequency oscillations and/or power fluctuations in high-frequency oscillations following the periodicity of the beat, plus its multiples or divisors. Critically, the stability of the connection between two or more active neural oscillations, i.e., the "resistance" to external perturbations, depends on the ratio of their periods (e.g., 1:1, 2:1, 2:3). Small integer ratios typically confer greater stability. This may explain the perceptual advantage for integer-ratio stimuli over more complex metrical patterns (Large and Kolen, 1995). Other frameworks state that specific neurons or neural channels are tuned to particular durational intervals or tempi (Merchant et al., 2013; Bartolo et al., 2014).

ITERATED DRUMMING EXPERIMENTS: SMALL INTEGER RATIOS AS COGNITIVE ATTRACTORS

Recent behavioral research investigated human priors for durations in rhythmic patterns (Ravignani et al., 2016, 2018; Jacoby and McDermott, 2017). Participants were given drumming sequences to reproduce to the best of their ability. The patterns produced were presented to the same or a new participant in an iterative procedure. Strikingly, "first-generation" participants were given completely random patterns, and "last-generation" participants produced rhythms exhibiting small integer ratios, in line with previous work on e.g., bimanual tapping (Peper et al., 1991, 1995a,b; Peper and Beek, 1998).

Specifically, participants were presented with sequences of IOIs sampled from a uniform distribution U (e.g., **Figure 1B**). As the patterns were transmitted through "chains of reproductions," (Ravignani et al., 2016, 2018; Jacoby and McDermott, 2017), distribution U converged toward a distribution D: a human observer's posterior distribution of IOIs (e.g., **Figure 1A**). This distribution is multimodal, and the modes are related by small integer ratios, a universal property of human musical cultures (Ravignani et al., 2016; Jacoby and McDermott, 2017).

Here we aim to explain the distribution D via established psychophysical principles, none of which explicitly entail small-integer ratios. In other words, is the integer ratio bias a perceptual primitive in itself, or might it arise from the interaction of more
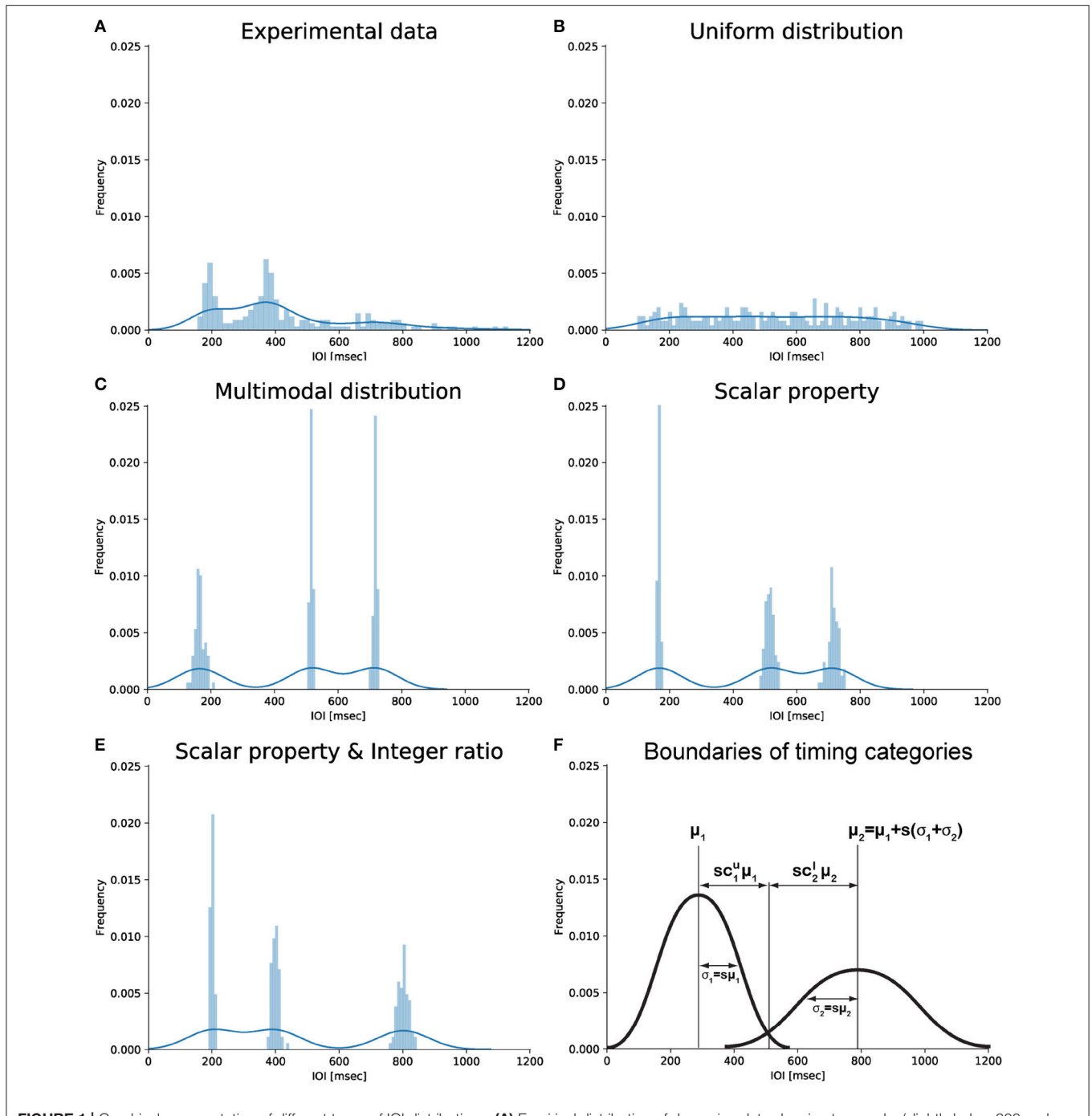
FIGURE 1 | Graphical representation of different types of IOI distributions. **(A)** Empirical distribution of drumming data showing two peaks (slightly below 200 and 400 ms) consistent with the notion of integer ratio categories. Data from the last experimental generation of chain 2 in Ravignani et al. (2016). **(B)** Uniform distribution from 100 to 1,000 ms. **(C)** Multimodal distribution based on 3 randomly chosen centroids without further assumptions. **(D)** Multimodal distribution around the same 3 centroids assuming the scalar timing property. **(E)** Multimodal distribution assuming the scalar timing property and showing small integer ratios. Data in panels **(B–E)** are simulated; they were randomly sampled from several normal distributions, with total sample size as in **(A)**. **(F)** Schematic representation of potential parameters linking scalar timing and small integer ratios. Panel **(F)** was produced without simulated or experimental data. Notice how the x-coordinate of the intersection point between the two Gaussians can be parameterized as to $\mu_1 + sc_1^u\mu_1$ (first Gaussian) and $\mu_2 - sc_2^l\mu_2$ (second Gaussian). For more than two Gaussians, the intersection can be parameterized as $\mu_k + sc_k^u\mu_k$ (first Gaussian) and $\mu_{k+1} - sc_{k+1}^l\mu_{k+1}$ (second Gaussian). This parameterization is used in the derivations below.

fundamental primitives? Jacoby and McDermott (2017) related a theoretically hypothesized prior with built-in integer ratios to an empirically estimated prior, showing that these were aligned.

Here, we investigate whether it is possible to derive a prior with similar properties by not building in the integer-ratio, but by combining empirically founded principles of timing with a

minimum of assumptions (and room for refinement by future testing).

PROBABILISTIC INFERENCE FOR INTERVAL RATIO CATEGORIES

Our concrete question is: Under which conditions will a distribution G show small-integer ratios, without having built these ratios into our model?

Without any assumptions, distribution G would equal the uniform IOI distribution U *in expectation*. In other words which results on basic mechanisms of rhythm perception and production allow us to turn U into G? Below, we make four assumptions based on psychophysical evidence and reduce the number of free parameters in the model drastically with little loss of generality. We begin by elaborating on previous formalizations to make relevant assumptions explicit and comparable.

ASSUMPTION 1: CATEGORICAL TIMING

An n-event rhythm defines a sequence of IOIs $\boldsymbol{d} = (d_1, \dots, d_{n-1})$ and of ratios $\boldsymbol{r} = (r_1, \dots, r_{n-2})$, such that $r_i = d_{i+1}/d_i$. Perception of a rhythm $\boldsymbol{r}$ induces a representation $\boldsymbol{z} = (z_1, \dots, z_{n-2})$, with a strong tendency to categorize. The vector $\boldsymbol{z}$ is a sequence of a small number of unique phenomenal interval-ratio *categories* that represent the observed data $\boldsymbol{r}$. More specifically, the notation $z_i = k$ identifies that interval ratio r_i is attributed to phenomenal category k (Ravignani et al., 2018). Whilst not used explicitly in our calculations, $\boldsymbol{z}$ formalizes the first key assumption: the processing of rhythmic sequences recruits a categorical interpretation of time intervals from a continuous stream of events (Clarke, 1987; Schulze, 1989; Desain and Honing, 2003). Behavioral evidence shows that also human motor timing is categorical: participants tapping produce IOI distributions with distinct peaks reflecting underlying durational *categories* (Collyer et al., 1994). This suggests that the distribution G can be approximated as a multimodal mixture of normal distributions (**Figure 1C**), rather than a uniform distribution (**Figure 1B**). A small number of durational categories naturally results in a small number of ratio categories. For the perception of a rhythmic sequence as a whole, we would argue that the perceived durations be transformed toward forming small ratios, as supported by iterated drumming experiments (Jacoby and McDermott, 2017), "ideally" into integer multiples of the smallest unit. Whilst categorical timing may appear to be a simplifying psychological concept (Schulze, 1989; Drake and Bertrand, 2001; Desain and Honing, 2003; ten Hoopen et al., 2006) based on behavioral observations, it may not be that far off neural reality. The notion of durational categories relate to basic durational tuning properties of premotor neurons recorded in non-human primates (Merchant et al., 2013). For instance, categories can be mapped to interval tuning in the premotor neurons of monkeys performing a synchronization continuation task (Merchant et al., 2013). Here, the distribution of preferred intervals could be viewed as a prior, although this distribution is multimodal, rather than bimodal as in Merchant et al. (2013). In addition, human

neuroimaging work showed specific activation patterns for the perceptual processing of integer interval ratios (Sakai et al., 1999). Moreover, sequences of small integer ratios may induce a metrical beat by the hierarchical organization of periodicity at two or more levels, i.e., the occrurence of an accent at a multiple small integer of the shortest time unit at the next higher level (Povel and Essens, 1985). Metrical structure is thus a higher, multi-level demonstration of the psychological prior toward small-integer ratios, that affords accurate reproduction (Povel and Essens, 1985). Moreover, the perceptual timing of rhythms with such a metrical beat is more accurate, their subjective percept "catchier" and their recognition more robust against temporal scaling, i.e., speeding up or slowing down the tempo, as the pattern is processed as one coherent whole rather than a series of time intervals, in contrast to rhythms that feature small integer ratios but no metrical beat (Grube and Griffiths, 2009).

ASSUMPTION 2: BAYESIAN INFERENCE OVER GAUSSIAN CATEGORIES

A general assumption in rhythm research is that perceptual timing can be described as a process combining prior beliefs with sensory input. One way to capture this mathematically is to model time perception as Bayesian inference (Jazayeri and Shadlen, 2010; Cicchini et al., 2012; Merchant et al., 2013; Pérez and Merchant, 2018). Whilst our analysis relies on the nature of the prior rather than how it is deployed during perceptual interpretation, taking a Bayesian viewpoint is useful. It lets us express a prior distribution as an inductive bias (Thompson et al., 2016) and has been successfully applied in previous models of time interval estimation (e.g., Jazayeri and Shadlen, 2010; Cicchini et al., 2012). Employing Bayesian inference, we can characterize participant behavior as attributing a categorical representation to interval ratio r_i according to the distribution $p(z_i = k \mid r_i) \propto p(r_i \mid z_i = k) p(z_i = k)$. Our focus is the prior distribution over categories, $p(z_i = k)$, equivalently G. Alternatively, it would be possible to model learners' assumptions about a likelihood distribution as a source of bias (e.g., Jazayeri and Shadlen, 2010; Cicchini et al., 2012).

Jacoby and McDermott (2017) recently modeled *n*-interval rhythms as single points in the *n-1* dimensional simplex, and formulated a multivariate-mixture prior over this space, assuming Gaussian models to underlie each of the mixtures. Namely, they formulated a multivariate $p(\boldsymbol{z})$ directly. Our approach to the prior is closely related. Like Jacoby and McDermott (2017), we express the prior as a mixture of Gaussian components. However, our formulation treats an *n*-interval rhythm as a set of *n-1* independent samples from a *univariate multimodal distribution*, rather than a single multivariate sample. The two approaches essentially represent minor variants of the model for covariance of interval ratio categories. The assumption that the distribution $p(\boldsymbol{z})$ has a Gaussian form should be tested in

future work, but is in line with existing work and a fair first approximation.

We write the prior as a K-dimensional Gaussian mixture of interval ratio categories, and the data likelihood as i.i.d. Gaussian underlying these categories, such that the marginal distribution of interval ratios has the form:

$$p(r) = G(r) = \prod_{i=1}^{n-1} \sum_{k=1}^{K} \varphi_k N(d_i; \mu_k, \sigma_k) \qquad (1)$$

Here, the prior assigns to each Gaussian $k = 1, ..., K$ a weight in the mixture, φ_k, which determines its relative prominence as a category; a category mean μ_k, which specifies the expected interval ratio underlying this category; and a category variance σ_k. The assumption we make is that weights are constant: $\varphi_k = K^{-1}$ (corresponding to an equal number of observations in the Gaussians in **Figures 1C–E**). Whilst we hope to examine this assumption empirically in the future, we proceed under the most neutral assumption: no interval-ratio category is privileged.

ASSUMPTION 3: A SMALL NUMBER OF SUB-SECOND CATEGORIES

Assuming that our indexing of categories under the prior is strictly ordered by the category means, such that $\mu_j < \mu_k \Leftrightarrow j < k$, we can immediately express our second empirical constraint on distribution G: only few categories exist (Desain and Honing, 2003; Motz et al., 2013; Ravignani et al., 2016, 2018). K is naturally limited by our approach to only model components for *small* integer ratios, and these are limited in number. Furthermore, we bound the range of category means μ_k from 200 ms (London, 2004, p. 35) to 1,000 ms (Shaffer, 1983; Desain and Honing, 2003; Buhusi and Meck, 2005). This constraint limits K to the largest number of categories such that no category mean exceeds 1,000 ms:

$$K = argmax_k \, \mu_k \, s.t. \, \mu_k \leq 1000 \text{ for } k = 1, \ldots, K. \qquad (2)$$

ASSUMPTION 4: SCALAR TIMING

So far, our assumptions constrain neither category means μ_k nor standard deviations σ_k. Our final, perhaps most central assumption is that timing exhibits *scalar properties* in the sub-second time range considered here (Gibbon, 1977; Matell and Meck, 2000). Scalar timing drastically reduces the number of free parameters describing distribution G, by expressing category variances as a function of category means. The standard deviation of each category σ_k equals the mean μ_k multiplied by a constant, dimensionless factor s (**Figure 1E**):

$$\sigma_k = s \, \mu_k. \qquad (3)$$

Previous empirical reports estimated s to approximate *0.025* (Friberg and Sundberg, 1995; Madison and Merker, 2004).

LINKING CATEGORICAL PERCEPTION AND SCALAR TIMING: HOW CLOSE CAN WE GET TO INTEGER RATIO INTERVALS?

All four assumptions are empirically based and independent of each other. Now, G can be further characterized by the degree of overlap between Gaussians composing the mixture. To formalize this, we assume each category k to intersect with its adjacent neighbors $k-1$ and $k+1$ at a distance proportional to c_k^l and c_k^u away from its mean μ_k (**Figure 1F**), which is a constant proportion of the standard deviation σ_k. c_k^l and c_k^u parameterize the overlap between categories: they express how many standard deviations away from its mean μ_k the cluster k intersects the cluster $k+1$, and how many standard deviations away from its mean μ_{k+1} the cluster $k+1$ intersects the cluster k (**Figure 1F** shows an example for $k = 1,2$).

Combining this idea of a parameterized overlap with scalar properties, each cluster k extends from $\mu_k - sc_k^l \mu_k$ to $\mu_k + sc_k^u \mu_k$. Under these assumptions, the distance between the means of two adjacent distributions (**Figure 1F**) can be written as

$$\mu_{k+1} - \mu_k = sc_{k+1}^l \mu_{k+1} + sc_k^u \mu_k, \qquad (4)$$

and their ratio as

$$r_k = \mu_{k+1}/\mu_k. \qquad (5)$$

Substituting (5) into (4) provides

$$r_k \mu_k - \mu_k = sc_{k+1}^l r_k \mu_k + sc_k^u \mu_k, \qquad (6)$$

which can be simplified and rewritten as

$$r_k = (1 + sc_k^u)/(1 - sc_{k+1}^l). \qquad (7)$$

Equation (7) requires, to be well-defined, that its right side is positive, namely

$$0 < c_{k+1}^l < \frac{1}{s}. \qquad (8)$$

Operationally, the category means following from the constraints on G can be calculated using the recursion equation:

$$\mu_{k+1} = r_k \mu_k. \qquad (9)$$

The constraints structure the space of component Gaussians in the prior such that, by specifying μ_1, we can compute μ_k for all $k \leq K$ using Equation (9) (**Figure 1E**).

These quantitative tools enable the formulation of several questions. Given our *post-hoc* knowledge that the prior is characterized by categories centered at small integer ratios, do the constraints we laid out structure the prior such that integer-ratio clusters are predicted by setting μ_1 to the smallest possible integer ratio?

An alternative approach might be to assume that one ratio is e.g., ½, and ask whether our equations imply small integer
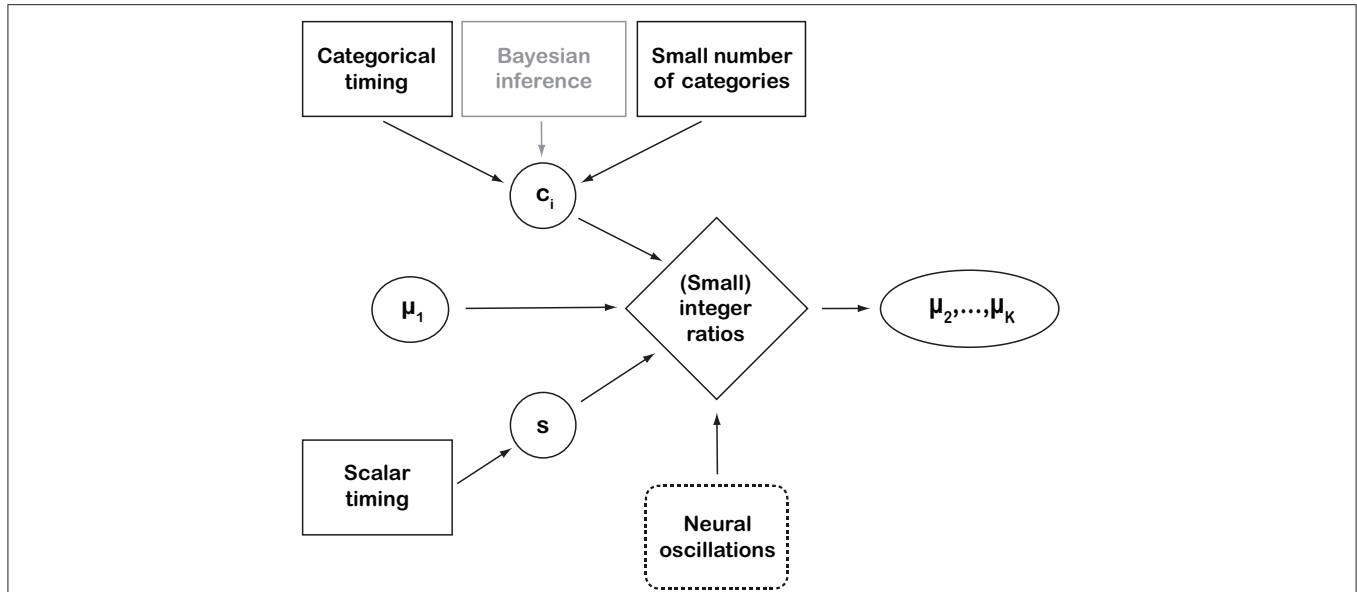
FIGURE 2 | Schematic representation of the perspective introduced by this paper. Black solid-line boxes represent empirically supported assumptions. "Bayesian inference" is outlined in gray to indicate that it is used here as a working assumption and conceptual framework, rather than an empirically supported assumption on cognitive processes (Shi et al., 2013). "Neural oscillations" are dashed because they represent observed neural process whose connection with the other behavioral concepts has not been proven (yet). The quantitative parameters are: category means μ_i, a scalar constant s, and c_i, which is the abbreviation of c_i^l and c_i^u, parameterizing the overlap between categories. The proposed way of representing rhythmic structure depends, among other factors, on the constancy of r_k (see main text). A deviation from this constancy would result in larger integer ratios, with the deviation accumulating over the categories when iterating equation (8). Empirical work (e.g., Ravignani et al., 2016; Jacoby and McDermott, 2017) has tried to operationalize the connection between the "mathematical perfection" of integer ratios and their empirical counterpart in a number of alternative ways. This perspective does not address how and when a real number is perceived as an integer ratio, leaving this as an empirical question for psychophysics research. In general, large integer ratios, and even irrational-number ratios, can be perceived as small integer ratios if close enough to one. For instance, $2^{7/12} \approx 1.498307$ is irrational (Coxeter, 1968) but close to 3/2. Virtually all pianos, today, employ this irrational number (1.498307) in their well-tempered tuning, which is "close enough" for human hearing to the integer ratio 3:2. At the same time, the "catchiness" of a rhythm also depends on small deviations from the integer ratios. For instance, delayed occurrences of expected beats even at varying levels of deviation from the underlying rhythms (together with the compensatory temporary speed-ups) are perceived as interesting, while a strictly regular rhythm will quickly appear dull.

ratios for the remaining cluster centers. More generally, do the constraints laid out impose an integer ratio structure on the prior without assuming an integer ratio for any of the clusters, simply by setting c_k in a certain way?

HOW DO c_k^u AND c_k^l RELATE TO μ_k ?

The x-coordinates for the intersection point, expressed as $\mu_k - sc_k^l \mu_k$ and $\mu_k + sc_k^u \mu_k$, can be substituted in the respective Gaussian probability density functions, equated to impose the condition of intersection on the y-axis (**Figure 1F**):

$$2 * log\left(s\mu_k\right) + \frac{\left(\left(\mu_k + sc_k^u \mu_k\right) - \mu_k\right)^2}{s^2 \mu_k^2}$$

$$= 2 * log\left(s\mu_{k+1}\right) + \frac{\left(\left(\mu_{k+1} - sc_{k+1}^l \mu_{k+1}\right) - \mu_{k+1}\right)^2}{s^2 \mu_{k+1}^2} \quad (10)$$

which simplifies as:

$$\left(c_k^u\right)^2 - \left(c_{k+1}^l\right)^2 = 2\,log(\mu_{k+1}/\mu_k). \quad (11)$$

Equation (11) means that the difference of squares between c's is proportional to the logarithm of the ratio of the two means.

To make an example with actual numbers, if one substitutes $\mu_k = \mu_1 = 100\ ms$ and $\mu_{k+1} = \mu_2 = 200\ ms$ in (11), the equation becomes $\left(c_k^u\right)^2 - \left(c_{k+1}^l\right)^2 = 2\,log(2)$. Hence $r_1 = \frac{\mu_{k+1}}{\mu_k} = 2$, $c_1^u \approx 2.5$ and $c_2^l \approx 2.2$ are two approximate solutions (among the infinite possible ones) of this particular example.

As the right side of Equation (11) is always strictly positive, c_k^u can never equal c_{k+1}^l. While this does not constitute a mathematical contradiction with our formulation (still leaving an infinite number of mathematically possible c's), it is admittedly difficult to interpret psychophysically.

SUGGESTED EXPERIMENTS: MODELING AND PSYCHOPHYSICS

Equations (7, 9) support a potential link between scalar timing and integer ratios, as they include the integer ratios r_k and the scalar constant s (**Figure 2**). These generative formulas can be implemented in computational simulations to explore the shape of the parameter space. Given specific values for parameters s, c_k^u and c_k^l, the equations will return a unique set of ratios: are

these small integer ratios? Likewise, given one single integer ratio μ_1, all other μ_k are determined by Equation (9): which values of μ_1 result in r being integer ratios and s, c_k^u and c_k^l being psychophysically plausible values?

The perspective we offer here creates the basis for expanding not only into theoretical but also empirical work on s, c_k^u and c_k^l. Experimental research can advance this approach by estimating s, c_k^u and c_k^l via Equation (7) or (11). Here, we treated the parameter s as an a priori known, one-valued constant ($s = 0.025$). To improve the model further, the variance of s might be estimated by replications of previous psychophysical experiments such as those by Friberg and Sundberg (1995) and Madison and Merker (2004). Values for c_k^u and c_k^l can be estimated from experiments testing the perception (and misattribution) of durational categories.

LIMITATIONS, DISCUSSION, AND CONCLUSIONS

We explore quantitative links between scalar timing and the human bias toward small integer ratios. The arguments we provide reduce the explanatory space to a few hypotheses. One possibility is that integer ratios are not a human cognitive primitive, but rather a simple by-product of other cognitive constraints, and their interaction.

Alternatively, the scalar timing framework might not be the most suitable one to explain the integer-ratio phenomenon of human rhythm. If one adopts oscillatory frameworks, integer ratios might simply arise from the oscillatory properties of brain activity, and so can scalar properties and categorical perception. Small integer ratios in particular would just reflect epiphenomena of harmonics of one oscillator or the interaction between two or more oscillators (Collyer et al., 1994; Strogatz, 2003; Buzsaki, 2006; Gupta, 2014; Merker, 2014; Gupta and Chen, 2016). Neural resonance to musical rhythm (Large, 2008), interval tuning

(Merchant et al., 2013; Bartolo et al., 2014), and population clocks (Crowe et al., 2014; Gouvêa et al., 2015; Bakhurin et al., 2016; Merchant and Averbeck, 2017) present alternative timing mechanisms, documented by in-vivo recordings of neural populations and compatible with the observed small integer bias.

In any case, scalar timing and oscillatory theories are simplifications, i.e., approximate descriptions derived from confined experimental set-ups. Neurally and behaviorally, the dissociation or compatibility between scalar timing and oscillatory theories is more complex than it may appear in higher level cognitive theories, and only detailed neural models will enable us to define the actual underlying mechanisms.

AUTHOR CONTRIBUTIONS

AR and BT conceived the idea and performed the mathematical derivations. All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Allman, M. J., and Meck, W. H. (2011). Pathophysiological distortions in time perception and timed performance. *Brain* 135, 656–677. doi: 10.1093/brain/awr210

Allman, M. J., Teki, S., Griffiths, T. D., and Meck, W. H. (2014). Properties of the internal clock: first- and second-order principles of subjective time. *Ann. Rev. Psychol.* 65, 743–771. doi: 10.1146/annurev-psych-010213-115117

Arnal, L. H., and Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends Cogn. Sci.* 16, 390–398. doi: 10.1016/j.tics.2012.05.003

Aubanel, V., Davis, C., and Kim, J. (2016). Exploring the role of brain oscillations in speech perception in noise: intelligibility of isochronously retimed speech. *Front. Hum. Neurosci.* 10:430. doi: 10.3389/fnhum.2016.00430

Bakhurin, K. I., Mac, V., Golshani, P., and Masmanidis, S. C. (2016). Temporal correlations among functionally specialized striatal neural ensembles in reward-conditioned mice. *J. Neurophysiol.* 115, 1521–1532. doi: 10.1152/jn.01037.2015

Bartolo, R., Prado, L., and Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *J. Neurosci.* 34, 3910–3923. doi: 10.1523/JNEUROSCI.2679-13.2014

Bizo, L. A., Chu, J. Y., Sanabria, F., and Killeen, P. R. (2006). The failure of Weber's law in time perception and production. *Behav. Process.* 71, 201–210. doi: 10.1016/j.beproc.2005.11.006

Buhusi, C. V., and Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* 6:755. doi: 10.1038/nrn1764

Buzsaki, G. (2006). *Rhythms of the Brain.* Oxford: Oxford University Press.

Celma-Miralles, A., de Menezes, R. F., and Toro, J. M. (2016). Look at the beat, feel the meter: top–down effects of meter induction on auditory and visual modalities. *Front. Hum. Neurosci.* 10:108. doi: 10.3389/fnhum.2016.00108

Church, R. M. (1999). Evaluation of quantitative theories of timing. *J. Exp. Anal. Behav.* 71, 253–256. doi: 10.1901/jeab.1999.71-253

Cicchini, G. M., Arrighi, R., Cecchetti, L., Giusti, M., and Burr, D. C. (2012). Optimal encoding of interval timing in expert percussionists. *J. Neurosci.* 32, 1056–1060. doi: 10.1523/JNEUROSCI.3411-11.2012

Clarke, E. F. (1987). "Categorical rhythm perception: an ecological perspective," in *Action and Perception in Rhythm and Music*, ed A. Gabrielsson (Stockholm: Royal Swedish Academy of Music), 19–33.

Collyer, C. E., Broadbent, H. A., and Church, R. M. (1994). Preferred rates of repetitive tapping and categorical time production. *Attent. Percept. Psychophys.* 55, 443–453. doi: 10.3758/BF03205301

Coxeter, H. S. M. (1968). Music and mathematics. *Math. Teach.* 61, 312–320.

Crowe, D. A., Zarco, W., Bartolo, R., and Merchant, H. (2014). Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *J. Neurosci.* 34, 11972–11983. doi: 10.1523/JNEUROSCI.2177-14.2014

Desain, P., and Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception* 32, 341–365. doi: 10.1068/p3370

Deutsch, D. (1986). A musical paradox. *Music Percept.* 3, 275–280. doi: 10.2307/40285337

Drake, C., and Bertrand, D. (2001). The quest for universals in temporal processing in music. *Ann. N. Y. Acad. Sci.* 930, 17–27. doi: 10.1111/j.1749-6632.2001.tb05722.x

Essens, P. J. (1986). Hierarchical organization of temporal patterns. *Percept. Psychophys.* 40, 69–73. doi: 10.3758/BF03208185

Essens, P. J., and Povel, D. (1985). Metrical and nonmetrical representations of temporal patterns. *Percept. Psychophys.* 37, 1–7. doi: 10.3758/BF03207132

Fraisse, P. (1982). "Rhythm and tempo," in *The Psychology of Music*, ed D. Deutsch (New York, NY: Academic Press), 149–180. doi: 10.1016/B978-0-12-213562-0.50010-3

Friberg, A., and Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *J. Acoust. Soc. Am.* 98, 2524–2531.

Getty, D. J. (1975). Discrimination of short temporal intervals: a comparison of two models. *Percept. Psychophys.* 18, 1–8.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* 84:279. doi: 10.1037/0033-295X.84.3.279

Gouvêa, T. S., Monteiro, T., Motiwala, A., Soares, S., Machens, C., and Paton, J. J. (2015). Striatal dynamics explain duration judgments. *Elife* 4:e11386. doi: 10.7554/eLife.11386

Grondin, S. (2001). From physical time to the first and second moments of psychological time. *Psychol. Bull.* 127:22. doi: 10.1037/0033-2909.127.1.22

Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. *Attent. Percept. Psychophys.* 72, 561–582. doi: 10.3758/APP.72.3.561

Grube, M., and Griffiths, T. D. (2009). Metricality-enhanced temporal encoding and the subjective perception of rhythmic sequences. *Cortex* 45, 72–79. doi: 10.1016/j.cortex.2008.01.006

Gupta, D. S. (2014). Processing of sub- and supra-second intervals in the primate brain results from the calibration of neuronal oscillators via sensory, motor, and feedback processes. *Front. Psychol.* 5:816. doi: 10.3389/fpsyg.2014.00816

Gupta, D. S., and Chen, L. (2016). Brain oscillations in perception, timing and action. *Curr. Opin. Behav. Sci.* 8, 161–166. doi: 10.1016/j.cobeha.2016.02.021

Ivry, R. B., and Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends Cogn. Sci.* 12, 273–280. doi: 10.1016/j.tics.2008.04.002

Jacoby, N., and McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Curr. Biol.* 27, 359–370. doi: 10.1016/j.cub.2016.12.031

Jazayeri, M., and Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nat. Neurosci.* 13:1020. doi: 10.1038/nn.2590

Jones, M. R., and Yee, W. (1997). Sensitivity to time change: the role of context and skill. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 693–709. doi: 10.1037/0096-1523.23.3.693

Karmarkar, U. R., and Buonomano, D. V. (2007). Timing in the absence of clocks: encoding time in neural network states. *Neuron* 53, 427–438. doi: 10.1016/j.neuron.2007.01.006

Large, E. W. (2008). Resonating to musical rhythm: theory and experiment. *Psychol. Time* 189–232. doi: 10.1016/B978-0-08046-977-5.00006-5

Large, E. W., and Jones, M. R. (1999). The dynamics of attending: how we track time varying events. *Psychol. Rev.* 106, 119–159. doi: 10.1037/0033-295X.106.1.119

Large, E. W., and Kolen, J. (1995). Resonance and the perception of musical meter. *Connect. Sci.* 6, 177–208. doi: 10.1080/09540099408915723

London, J. (2004). *Hearing in Time.* New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195160819.001.0001

Madison, G., and Merker, B. (2004). Human sensorimotor tracking of continuous subliminal deviations from isochrony. *Neurosci. Lett.* 370, 69–73. doi: 10.1016/j.neulet.2004.07.094

Matell, M. S., and Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *Bioessays* 22, 94–103. doi: 10.1002/(SICI)1521-1878(200001)22:1<94::AID-BIES14>3.0.CO;2-E

Mauk, M. D., and Buonomano, D. V. (2004). The neural basis of temporal processing. *Annu. Rev. Neurosci.* 27, 307–340. doi: 10.1146/annurev.neuro.27.070203.144247

Meck, W. H. (1996). Neuropharmacology of timing and time perception. *Cogn. Brain Res.* 3, 227–242.

Merchant, H., and Averbeck, B. B. (2017). The computational and neural basis of rhythmic timing in medial premotor cortex. *J. Neurosci.* 37, 4552–4564. doi: 10.1523/JNEUROSCI.0367-17.2017

Merchant, H., Pérez, O., Zarco, W., and Gámez, J. (2013). Interval tuning in the primate medial premotor cortex as a general timing mechanism. *J. Neurosci.* 33, 9082–9096. doi: 10.1523/JNEUROSCI.5513-12.2013

Merker, B. (2014). Groove or swing as distributed rhythmic consonance: introducing the groove matrix. *Front. Hum. Neurosci.* 8:454. doi: 10.3389/fnhum.2014.00454

Motz, B. A., Erickson, M. A., and Hetrick, W. P. (2013). To the beat of your own drum: cortical regularization of non-integer ratio rhythms toward metrical patterns. *Brain Cogn.* 81, 329–336. doi: 10.1016/j.bandc.2013.01.005

Palmer, C., and Krumhansl, C. L. (1990). Mental representations for musical meter. *J. Exp. Psychol.* 16:728. doi: 10.1037/0096-1523.16.4.728

Patel, A. D., Iversen, J. R., Chen, Y., and Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Exp. Brain Res.* 163, 226–238. doi: 10.1007/s00221-004-2159-8

Peper, C. E., and Beek, P. J. (1998). Distinguishing between the effects of frequency and amplitude on interlimb coupling in tapping a 2: 3 polyrhythm. *Exp. Brain Res.* 118, 78–92. doi: 10.1007/s002210050257

Peper, C. E., Beek, P. J., and van Wieringen, P. C. (1995a). Frequency-induced phase transitions in bimanual tapping. *Biol. Cybernet.* 73, 301–309.

Peper, C. E., Beek, P. J., and Van Wieringen, P. C. W. (1991). "Bifurcations in polyrhythmic tapping: in search of Farey principles," in *Tutorials in Motor Neuroscience,* eds J. Requin and G. E. Stelmach (Dordrecht: Springer), 413–431.

Peper, C. L. E., Beek, P. J., and van Wieringen, P. C. (1995b). Coupling strength in tapping a 2: 3 polyrhythm. *Hum. Mov. Sci.* 14, 217–245.

Pérez, O., and Merchant, H. (2018). The synaptic properties of cells define the hallmarks of interval timing in a recurrent neural network. *J. Neurosci.* 38, 4186–4199. doi: 10.1523/JNEUROSCI.2651-17.2018

Pikovsky, A., Rosenblum, M., and Kurths, J. (2003). *Synchronization: A Universal Concept in Nonlinear Sciences*, Vol. 12. Cambridge: Cambridge University Press.

Polak, R., London, J., and Jacoby, N. (2016). Both isochronous and non-isochronous metrical subdivision afford precise and stable ensemble entrainment: a corpus study of malian jembe drumming. *Front. Neurosci.* 10:285. doi: 10.3389/fnins.2016.00285

Povel, D. J., and Essens, P. (1985). Perception of temporal patterns. *Music Percept.* 2, 411–440. doi: 10.2307/40285311

Ravignani, A., Delgado, T., and Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nat. Hum. Behav.* 1:0007. doi: 10.1038/s41562-016-0007

Ravignani, A., Thompson, B., Grossi, T., Delgado, T., and Kirby, S. (2018). Evolving building blocks of rhythm: how human cognition creates music via cultural transmission. *Ann. N.Y. Acad. Sci.* 1423, 176–187. doi: 10.1111/nyas.13610

Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Tamada, T., Iwata, N. K., et al. (1999). Neural representation of a rhythm depends on its interval ratio. *J. Neurosci.* 19, 10074–10081. doi: 10.1523/JNEUROSCI.19-22-10074.1999

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. U.S.A.* 112, 8987–8992. doi: 10.1073/pnas.1414495112

Schulze, H. H. (1989). Categorical perception of rhythmic patterns. *Psychol. Res.* 51, 10–15. doi: 10.1007/BF00309270

Shaffer, L. H. (1983). Timing in musical performance. *Ann. N. Y. Acad. Sci.* 423, 420–428.

Shi, Z., Church, R. M., and Meck, W. H. (2013). Bayesian optimization of time perception. *Trends in Cognitive Sciences*, 17, 556–564. doi: 10.1016/j.tics.2013.09.009

Strogatz, S. H. (2003). *Sync: The Emerging Science of Spontaneous Order*. New York, NY: Hyperion Books.

ten Hoopen, G., Sasaki, T., Nakajima, Y., Remijn, G., Massier, B., Rhebergen, K. S., et al. (2006). Time-shrinking and categorical temporal ratio perception: evidence for a 1: 1 temporal category. *Music Percept.* 24, 1–22. doi: 10.1525/mp.2006.24.1.1

Thompson, B., Kirby, S., and Smith, K. (2016). Culture shapes the evolution of cognition. *Proc. Natl. Acad. Sci. U.S.A.* 113, 4530–4535. doi: 10.1073/pnas.1523631113

Toussaint, G. T. (2013). *The Geometry of Musical Rhythm: What Makes a Rhythm "Good"?* New York, NY: Chapman and Hall; CRC Press.

Wearden, J. H. (1991). Do humans possess an internal clock with scalar timing properties? *Learn. Motiv.* 22, 59–83. doi: 10.1016/0023-9690(91)90017-3

Wing, A. M., and Kristofferson, A. B. (1973a). Response delays and the timing of discrete motor responses. *Atten. Percept. Psychophys.* 14, 5–12.

Wing, A. M., and Kristofferson, A. B. (1973b). The timing of interresponse intervals. *Atten. Percept. Psychophys.* 13, 455–460.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.