

## RHYTHMIC PROTOTYPES ACROSS CULTURES: A COMPARATIVE STUDY OF TAPPING SYNCHRONIZATION

---

RAINER POLAK

*Max Planck Institute for Empirical Aesthetics,  
Frankfurt, Germany*

NORI JACOBY

*Columbia University*

TIMO FISCHINGER

*Max Planck Institute for Empirical Aesthetics,  
Frankfurt, Germany*

DANIEL GOLDBERG

*University of Connecticut*

ANDRE HOLZAPFEL

*KTH Royal Institute of Technology, Stockholm,  
Sweden*

JUSTIN LONDON

*Carleton College*

**IT HAS LONG BEEN ASSUMED THAT RHYTHM** cognition builds on perceptual categories tied to prototypes defined by small-integer ratios, such as 1:1 and 2:1. This study aims to evaluate the relative contributions of both generic constraints and selected cultural particularities in shaping rhythmic prototypes. We experimentally tested musicians' synchronization (finger tapping) with simple periodic rhythms at two different tempi with participants in Mali, Bulgaria, and Germany. We found support both for the classic assumption that 1:1 and 2:1 prototypes are widespread across cultures and for culture-dependent prototypes characterized by more complex ratios such as 3:2 and 4:3. Our findings suggest that music-cultural environments specify links between music performance patterns and perceptual prototypes.

*Received August 3, 2017, accepted April 7, 2018.*

**Key words:** culture, universals, cross-cultural comparison, categorical rhythm perception, rhythmic prototype

---

**T**HIS PAPER FOCUSES ON PROTOTYPES IN rhythm perception and production. While the timing of musical events is infinitely variative, humans perceive but a few rhythmic categories, which serve as the building blocks for more extended rhythmic structures. It has been broadly argued that rhythmic categories and their metric relationships are closely tied to prototypical durational patterns that lie near small-integer ratios, particularly the simplest ratios such as 1:1 and 2:1—the former being an isochronous series of pulses or onsets, and the latter a long-short pattern of durations/interonset intervals (among many others, see Clarke, 1985, 1987; Desain & Honing, 2003; Drake, 1993; Drake & Bertrand, 2001; Essens & Povel, 1985; Fraisse, 1956, 1982; Large, 2008; Large & Jones, 1999; Large & Palmer, 2002; Large & Snyder, 2009; Lerdahl & Jackendoff, 1983; Longuet-Higgins & Lee, 1982; Madison & Merker, 2002; Merker, Madison, & Eckerdal, 2009; Povel, 1981, 1984; Ravignani, Delgado, & Kirby, 2016; Ravignani & Madison, 2017).

The degree to which perceptual categories are either learned or innate is often unclear, though it is well established that perceptual learning for these categories plays an important role in some domains (Goldstone, Leeuw, & Landy, 2015). The perception of phonetic categories in speech is a particularly well-researched case in point; phonemic perception varies with the linguistic context (Bradlow, Akahane-Yamada, Pisoni, & Tohkura, 1999; Goldstone, 1998; Goldstone & Hendrickson, 2010; Lively & Pisoni, 1997; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994; Repp, 1984). By contrast, it is often implicitly and sometimes explicitly assumed that the rhythmic prototypes tied to the simplest integer ratios (1:1 and 2:1) will be identical across cultures (e.g., Drake & Bertrand, 2001; Epstein, 1985; Povel, 1981). For example, it has been argued that rhythms with small integer proportions are universally privileged as a direct consequence of mechanisms of resonance of neural oscillations in the human brain (Large, 2008; Large & Kolen, 1994).

However, humans' learning of rhythmic categories is a dynamic and ongoing process, one that may occur during the span of a single laboratory experiment as well as an entire lifetime. For example, in a study of the

identification of periodic rhythms comprised of three intervals, Schulze (1989) found that, after a training session, musicians were able to identify a large number of rhythmic categories. Likewise, Clarke (1987) and Desain and Honing (2003) showed that category boundaries are open to influence from metric priming. Honing noted that, “it is the most frequently heard performance of a rhythm—rather than the canonical or integer-related version of it, as notated in a score—that might function as a reference or category [...] which, in turn, is dependent on memory, expectation, and the ways in which we have been exposed to music in the course of our lives” (Honing, 2013, p. 380). Thus, effects of long-term, culture-specific exposure and experience upon rhythmic category learning seem likely. Indeed, a growing body of cross-cultural research finds both linguistic and music-cultural backgrounds to influence rhythm perception and processing (Cameron, Bentley, & Grahn, 2015; Drake & El Heni, 2003; Iversen, Patel, & Ohgushi, 2008; Stobart & Cross, 2000; Toiviainen & Eerola, 2003; Will, 2017). In particular, enculturated familiarity with music showing mildly complex rhythmic ratios such as 2:2:3 can positively override the potential difficulty of processing such rhythms.<sup>1</sup> These rhythms are relatively difficult for North Americans to perceive (Repp, London, & Keller, 2005; Snyder, Hannon, Large, & Christiansen, 2006), but not for participants with Balkan, Turkish, or Indian backgrounds (Hannon, Soley, & Ullal-Gupta, 2012; Kalender, Trehub, & Schellenberg, 2013; Ullal-Gupta, Hannon, & Snyder, 2014), as the music in these countries more prominently features such rhythmic patterns, nor for very young North American infants, suggesting that the adult North Americans’ bias towards isochronous meters reflect their specific enculturation

<sup>1</sup> Here and in the following, we comply with terminological usage in the literature on rhythm perception in referring to only the very simplest of small-integer ratios (1:1, 2:1, and 3:1) as “simple.” By contrast, we speak of 3:2, 4:3, or larger integer ratios as “(mildly) complex” rhythms. This may appear mathematically dubious; for instance, the ratio of 3:2 consists of very small integers and is mathematically simple. However, its dividend is not an integer multiple of the divisor; in other words, the quotient is not also an integer. This involves a somewhat higher degree of rhythmic complexity. For instance, 1:1, 2:1, and 3:1 are simple in that their greatest common divisor ( $\text{gcd}=1$ ) can be used as a unit of measure for interpolation-based coordination of the two durations ( $1:1 = xx$ ,  $2:1 = x.x$ ,  $3:1 = x.x$ ). This property is advantageous, as it affords a metric subdivision benefit for rhythm perception (London, 2012; Martens, 2011; Repp, 2003, 2008). By contrast, a gcd-based mental plotting of a 3:2 pattern would require quintuplet subdivision ( $x.x$ ). The  $\text{gcd} (= 1)$  here is not directly suggested by any of the two sounding intervals ( $= 2, 3$ ) and requires more interpolations ( $= 3$ ) relative to the realized onsets ( $= 2$ ); furthermore, it is more rapid relative to the rhythmic pattern. These features can be assumed to result in a drastic decrease of potential subdivision benefit for processing this rhythm.

rather than a biological predisposition (Hannon & Trehub, 2005a). Infants and young children can easily learn to perceive these kinds of complex rhythms as a result of passive exposure; such learning is achieved less rapidly and fluently by older children and adults (Hannon & Trehub, 2005b; Hannon, Vanden Bosch der Nederlanden, & Tichko, 2012).

In sum, there is evidence to justify the assumption that the characteristics of specific musical enculturation can play a strong role in the cognition of musical rhythm. Ethnomusicological research has shown both cross-cultural trends toward statistical universals (Brown & Jordania, 2013) as well as abundant diversity in human rhythmic practices. One example of a rhythmic feature put forth as a statistical universal is isochronous metric beats: While pulse streams of beats with 1:1 proportions are not a feature of all music, Savage and colleagues argue that they feature in over 86% of the recordings accompanying the largest ethnomusicological encyclopedia of the world’s music published to date (Savage, Brown, Sakai, & Currie, 2015). However, this wide distribution of isochrony does not preclude other principles of rhythmic organization. For instance, metric cycles may contain non-isochronous beats, which is often the case in musical styles from Scandinavia, the Balkans, Turkey, the Near East, and Australia, among others (see Bates, 2011; Brăiloiu, 1951/1984; Cler, 1994; During, 1997; Goldberg, 2015, 2017; Haugen, 2016; Holzapfel, 2015; Johansson, 2009, 2017; Kivifte, 2007; Marcus, 2007; Will, 2011). Furthermore, isochronous metric beats can contain subdivisions related by complex integer or non-integer ratios; such “swung” rhythms feature prominently in the Sahel and Savannah zones of sub-Saharan Africa (Kubik, 2010, pp. 50–52; Polak, 2010; Polak & London, 2014) and Afro-diasporic musical styles from North Africa and the Americas (see Benadon, 2006; Butterfield, 2011; Gerischer, 2003, 2006; Jankowsky, 2013; Naveda, Gouyon, Guedes, & Leman, 2011).

#### TEMPO AND METRIC LEVELS

Previous research in the cross-cultural variation of rhythm perception has tended to focus on moderate tempos characteristic of rhythm at the metric level of the beat. In particular, this research has capitalized on the prominence of both isochronous (“simple”) and non-isochronous (“complex”) beat cycles in some culture-geographic areas (Balkans, Turkey, the Near East, and India) and the relative lack of non-isochronous meters in others (Western Europe, USA) (see Yates, Justus, Atalay, Mert, & Trehub, 2017). The present study expands that line of research by focusing also on the relatively fast tempos that occur at the

rhythmic surface level in many musical styles, often corresponding to the metric level of beat subdivision.

Our study is specifically motivated by recent research in the performance and perception of non-isochronous structures of beat subdivisions in Malian drum ensemble music for social dancing. This repertoire features rhythms that involve complex-ratio beat subdivisions in the range of 5:4, 4:3 and 3:2. These are highly stable across a wide range of rhythms, tempo changes, instruments, performances, and performers (Polak, 2010; Polak & London, 2014) and afford precise ensemble entrainment (Polak, Jacoby, & London, 2016). Neuhoff, Polak, and Fischinger (2017) found that timing patterns based on these complex ratios were well discriminated and aesthetically preferred by Malian expert musicians and dancers over both isochronous versions and complex ratios that were foreign to their repertoire. Here, we ask whether musicians' perception and production prototypes co-vary with the prominent occurrence of corresponding rhythms characteristic of their music-cultural environments. In particular, we study whether Malian musicians who have learned to listen and perform in the context of a musical culture that prominently features fast complex ratio two-element rhythms may also have developed a corresponding cognitive prototype active in the same fast tempo range. We examine this issue by comparing the performance of expert musicians from Mali with that of other experienced musicians from Bulgaria and Germany in a simple rhythm synchronization task, where participants tap along with all events in periodic two-element rhythm patterns.

#### MUSICIANSHIP

Previous tapping studies have rarely used intervals shorter than 250 ms, assuming that nonmusician participants usually cannot tap much faster than that without considerable interference from motor constraints (Povel, 1981; Repp, 2003). By contrast, the threshold for the fastest metric subdivision rate in music performance and perception has been suggested to go down to 125 ms (Repp, 2003), 100 ms (London, 2002), or even 80 ms (Polak, 2017). Basic synchronization or synchronization-continuation tapping paradigms thus cannot easily cover the fastest event activity rates found in music when testing the general population. Since this paper is testing a hypothesis about perception and production of rhythm at fast tempi (our experiment included stimuli with intervals as short as 167 ms), we focused mainly on musicians; additional results for nonmusicians at a slower tempo are reported in the Appendix.

A cross-cultural definition of "musician" can be problematic. While in some social contexts, a specific cohort

of the population may identify as musicians, in other cultures such distinctions may not exist. For instance, everybody may equally engage in performance, or what we here consider to be music may be subsumed as just one dimension of a multi-modal performance or as part of a ritual (Trehub, Becker, & Morley, 2015). In cases where identifiable groups of musicians exist, there are still differences: they may receive institutionalized training and work professionally in some cultures, while in others they do not. The commonly used measure of "years of training" here would be irrelevant.

In our study, we do not aim at a cross-culturally tenable definition of "musicianship," but rather use the concept as a heuristic for recruitment strategies. We targeted groups whose individual members: a) are typically identified and self-identify as specialized players of instruments in their respective social environments, b) are active as performers in public settings, and c) make major parts of their living as (semi-)professional players or are in the final stage of training for this profession. This operational definition covers European professional musicians and music conservatory students as well as professional drummers in urban Mali.

#### CULTURE

The empirical description of the structure and extent of cross-cultural variation is an indispensable first step toward an understanding of the relationship between universal constraints and cultural contingency in human musical behavior (Stevens, 2012). What we can empirically observe and should want to report is the extent of variation, without discarding strong degrees of variation as irrelevant from a "universalist" perspective or ignoring strong degrees of uniformity as uninteresting from a "relativist" perspective.

In the context of this cross-cultural research program, we conceive of culture as the human disposition toward social learning and the dynamic convergence of behavior within social fields as well as diversification across social fields. Social fields are spaces of increased likelihood of interaction, characterized by individuals' positions and potentials for action (Bourdieu, 1984). Cultural anthropologists have emphasized the multitude and difference of competing definitions of culture (see Kroeber & Kluckhohn, 1952), many of which propose to focus on binary aspects such as structure *or* content, overt behavior *or* meaning, etc. Our focus on social learning instead has the advantage of being inclusive, as it potentially acknowledges any aspect of behavior, including perception and cognition. This breadth has characterized many prominent definitions ever since Edward B. Tylor's seminal formulation: "Culture

or civilization, taken in its wide ethnographic sense, is that complex whole which includes knowledge, belief, art, morals, law, custom, and any other capabilities and habits acquired by man as a member of society” (Tylor 1871/1920, p. 1). Mechanisms of social learning and cultural transmission leading toward both relative continuity within social fields and diversity across fields have been theorized and modeled in studies of cultural evolution over the past decades (Boyd & Richerson, 2005; Henrich & Tennie, 2017; Richerson & Boyd, 2005; Tennie, Call, & Tomasello, 2009; Tomasello, 1999; Tomasello, Kruger, & Ratner, 1993). However, ethnographic research has amply demonstrated that “cultures” are fuzzy, malleable, changing, heterogeneous entities, often incongruent with ethnic, linguistic, political, or other units. Relatively discrete population units (ethnic groups, nations, states) often are *not* congruent with “a culture.” Of course, such social units *can* construct cultural difference and emphasize group boundaries, yet these are dynamic, often emergent outcomes of socio-cultural processes, not the stable essentials that, for instance, the concept of race would suggest (Amselle & M’Bokolo, 2009; Barth, 1998; Bashkow, 2004; Bourdieu, 1984; Brumann, 1999; Goody, 1992; Gupta & Ferguson, 1992; Kahn, 1989; Kuper, 1999/2003; Lentz, 2016).

Remaining mindful of these complexities of culture, we ran experiments in Bulgaria, Germany, and Mali, and assumed country of residence to represent only a rough proxy for partly overlapping yet partly non-overlapping social fields and music-cultural environments. Some aspects of these cultural contexts are probably relatively local, regional, or national, and thus potentially distinct, while others are transnational or global. It is of course impossible to quantify the degrees and details of socio-cultural overlap amongst the three countries considered. However, we can assume that degrees of geographical and social distance do vary and play a role for relative cultural distance (Elfenbein & Ambady, 2003). In an effort to acknowledge potential cultural variation not only across but also within countries, we included two participant groups differentiated by stylistic orientation in the case of Bulgaria, namely, a group of musicians who specialize in traditional Bulgarian music and a group of musicians trained in Western classical music.

#### CATEGORICAL PERCEPTION OF TWO-ELEMENT RHYTHMS

Categorical perception is a non-continuous discrimination mechanism that involves the warping of percepts in phenomena that physically vary along a sensory continuum. It groups together the instances of a certain segment of that continuum into discrete equivalence classes, or categories. The perceived difference between instances

within categories is perceptually compressed, whereas it is increased at the border between two categories (Harnad, 1990; Repp, 1984; Goldstone & Hendrickson, 2010).

Categorical perception has traditionally been studied using discrimination paradigms, typically a two-alternative forced-choice task (2AFC) task, which was introduced to the study of categorical rhythm perception by Clarke (1987). Clarke (1987) and Desain and Honing (2003) used another experimental paradigm when they asked musician participants to identify sounding rhythmic patterns with nominal durational values taken from Western musical notation. Reliance on notation has the disadvantage of limiting recruitment of participants to a single cultural tradition. It would be impossible to apply this method to the majority of worldwide music-cultural traditions, which do not typically use graphic symbols for music composition, pedagogy, and performance.

While both of these methods have proved highly productive, they also may risk overemphasizing the functional importance of boundaries and underestimating the core function of real-life categorical perception (Repp, 1984, pp. 319–320), which is to reliably and fluently identify “contents” (things, concepts, events, etc.), or at least to reduce uncertainty about what these contents are by perceiving them as instantiations of categories.

The characterization of perceptual categories often combines information about their borders with other information that helps differentiate among instances within categories, such as when we perceive slightly different shades of a color, a nuanced difference in the pitch inflection or intonation of a specific note, or two expressively timed eighth notes as variations of a rhythmic category (cf. Honing, 2013, p. 375). A core notion here is prototypicality, understood as the mechanism whereby a focal, particularly representative spot within the space of a category serves as a reference point (Rosch, 1975). Research in linguistic perception shows that prototypical members of a phonetic category are processed more efficiently than less typical ones, and that the latter are judged with reference to their distance from the former (Lively, Logan, & Pisoni, 1993; Samuel, 1982). Moreover, a prototype of a phonetic category functions like a “perceptual magnet,” which lowers discrimination accuracy between members of a category that approximate the prototype (Kuhl, 1991; Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992). Percepts are “pulled” toward perceptual prototypes because these correspond to the statistically most likely occurrence of the category, which the perceiver probabilistically estimates from his or her experience. Within-category percepts thus are biased toward the prototype (Clarke, 1987). A prototype with this biasing effect has been

referred to as an “attractor” in rhythm research (Repp, London, & Keller, 2012).

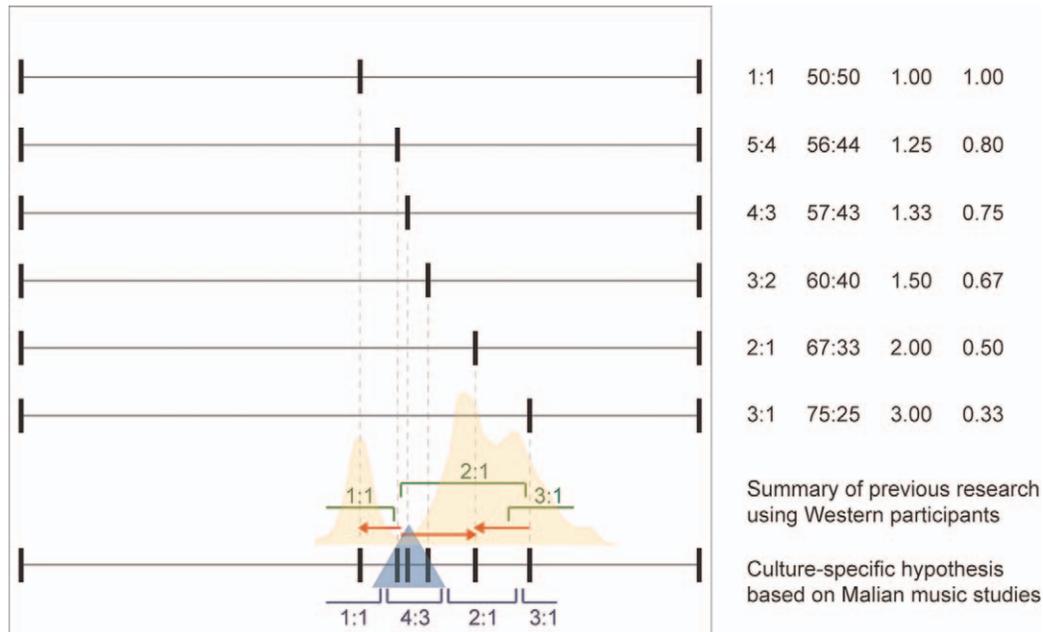
The concept of attractor/prototype as an aspect of categorical rhythm perception overlaps to some degree with the Bayesian concept of perceptual priors (Jacoby & McDermott, 2017; Sadakata, Desain, & Honing, 2006). It has been generally assumed that sensory systems try to extract true sensory information from a noisy measurement. Bayesian perception extends classical psychophysics by proposing that this process can be modeled as probabilistic inference, where both a priori knowledge about the external world and sensory measurement are taken into account. According to this view, observers are biased toward inferring that the perceived stimulus originates from a source in the real world that is more likely a priori. Therefore, perceptual attractors are simply highly probable points in a perceptual space, and their “attraction” is a consequence of the aforementioned inferential process. This conceptualization also provides a Bayesian interpretation for the “perceptual magnet” effect (Feldman, Griffiths, & Morgan, 2009).

Finger-tapping experiments whereby participants either tap their finger(s) along to a rhythmic “target” stimulus or continue tapping after an initial synchronization phase have been used extensively to study the production and perception of simple rhythms. Reliable reproduction is assumed to indicate that a stimulus ratio closely matches a perceptual rhythmic prototype. Consistent distortion away from the pacing stimulus’s rhythmic ratio, by contrast, is assumed to indicate that it does not match any mental representation, and that the observed distortion is due to the attraction of a categorically related rhythmic prototype that is physically different yet perceived as embracing the target stimulus. Acknowledging that the experimental paradigm of tapping relies on both perception and production, previous research made substantial efforts to quantify the relative contribution of these components. One approach was to test whether changing the type of production used within the experiment would significantly alter measured production biases. The idea here was that if these biases were determined by production constraints, then a change in the mode of production (for example by replacing unimanual tapping with bimanual tapping) would have a substantial effect on the results. Summers, Bell, and Burns (1989) tested tapping with one finger, two fingers, and one finger of opposite hands; Repp et al. (2012) compared unimanual and bimanual tapping; and Jacoby and McDermott (2017) compared finger tapping to a condition in which the reproductions were verbal (nonsense syllables). These experiments repeatedly demonstrated that the mode of production

has almost no effect on the average production biases, consistent with the idea that biases are predominately perceptual. In addition, in a meta-analysis, Sadakata et al. (2006) compared experiments that involved perception, production, or both perception and production. They concluded that these experimental results are highly consistent with a unified explanation under which perception and production are determined predominantly by identical mechanisms. More directly, Jacoby & McDermott (2017) showed that the results of a perception-only categorical discrimination experiment (as proposed by Clarke, 1987) are almost fully predicted from a finger tapping experiment, suggesting that participants use similar perceptual categories, or perceptual priors, to perform both discrimination and reproduction tasks (see Jacoby & McDermott, 2017, Figure 5 and Experiment 5).

Rhythms comprised of two interonset intervals that are periodically repeated occur in many music performance contexts and are basic building blocks of more complex rhythmic patterns. Several studies using the tapping paradigm found a strong bias towards a prototypical 2:1 ratio for all uneven two-element rhythms: Ratios that are “sharper” than 2:1, such as 3:1, are consistently “softened” in the direction of 2:1, whereas softer ratios such as 3:2 are sharpened in the direction of 2:1 (Essens & Povel, 1985; Jacoby & McDermott, 2017; Povel, 1981; Repp et al., 2005, 2011, 2012; Semjen & Ivry, 2001; Snyder et al., 2006; Summers et al., 1989; Summers, Hawkins, & Mayers, 1986). These findings are in agreement with seminal research in categorical rhythm perception based on perception-only paradigms, which suggested a prominent role for only two overarching categories in rhythm perception, namely, even and uneven, with prototypes at 1:1 and 2:1 (Clarke, 1987; Drake, 1993; Fraisse, 1956, 1982).

The upper tier of Figure 1 displays a selection of small-integer ratios to illustrate the physically continuous space of periodic two-interval rhythms. The lower tier schematically depicts the above findings of tapping research. The red arrows indicate consistent distortions that suggest an overarching role for the category characterized by the 2:1 prototype (horizontal bracket). This finding was recently supported and differentiated by Jacoby and McDermott (2017, Figure S1), who had participants repeatedly tap stimuli which are created on the fly from their own previous reproductions; this iterative paradigm allows experimenters to elicit prototypes with great precision and reliability. As the sand-colored histogram shows, the 2:1 category is powerful in terms of its overarching extension (roughly identical with the catchment area of distortions in the conventional,



**FIGURE 1.** Upper tier: Selection of small-integer ratios illustrating various two-element rhythms at a given period. We specify ratios in four different formats (from left to right): small-integer ratio, percentages of the period, quotient of the long-short ratio, and quotient of short-long ratio. Lower tier: Schematic representation of the perceptual categorization of the space of two-element rhythms. In the upper part of the lower tier, red arrows roughly summarize the main biases found in classical tapping studies using “Western” listeners (e.g., Essens & Povel, 1985; Povel, 1981; Repp et al., 2005, 2011). The sand-colored histogram shows the distribution of participants’ stabilized responses in the recently developed iterative tapping paradigm (taken from Jacoby & MacDermott 2017, Figure S1). Horizontal green brackets schematically indicate categories suggested by this research, with prototypes at 1:1, near 2:1, and near 3:1. The bottom part of the lower tier indicates the hypothesis, based on research in music from Mali (Polak, 2010; Polak & London, 2014; Polak et al., 2016) of an additional category with a prototype at about 4:3.

non-iterative tapping paradigm) and separated from the 1:1 category by a deep and broad valley of very low distribution frequency. The “empty zone” in the valley is indicative of perceptual ambiguity, which is characteristic of the border zones between two neighboring categories (Clarke, 1987).

This figure also shows that the 2:1 category is slightly biased away from the mathematical integer ratio proportion towards the 1:1 category. Such “softening” of the 2:1 ratio has been argued to be a characteristic of systematic variations of durations in expressive music performance (Gabrielsson, Bengtsson, & Gabrielsson, 1983), for example in the phenomenon referred to as “swing timing” in jazz research (Benadon, 2006). Indeed, Repp et al. (2011, 2012) found that participants’ responses in a tapping experiment were slightly softening the 2:1 prototype. Note, however, that other studies showed no bias or even a bias in the reverse direction of a “sharpened” 2:1 (e.g., Povel, 1981; see review in Repp et al., 2012).

In summary, an impressive body of experimental research converges in suggesting the existence of only two powerful perceptual categories in the space of two-element rhythms, even and uneven (1:1 and

approximately 2:1). A third category for very uneven rhythms with a prototype in the proximity of 3:1 is perceptually less salient and only vaguely distinct from the 2:1 category; its prototype overlaps and is attracted by the 2:1 prototype. However, all of the relevant research used exclusively “Western” participants.

#### HYPOTHESES

If rhythm perception were determined by basic biological or psychological constraints that are similar for all humans (for example, by mechanisms such as neural oscillations in the brain), we would expect culturally universal rhythmic prototypes. According to this “universalist” hypothesis, other factors such as the context in which the rhythms are presented and the musicians’ culture-specific experience would play minor roles in determining the characteristics of rhythmic perceptual categories. Alternatively, if perceptual categories were fundamentally dependent on culture, we would expect specific cultural environments to correspond to specific categories—a “relativist” hypothesis.

To operationalize these extreme hypotheses, we focus on the possibility that rhythmic prototypes may occur

where previous research converged in indicating that there are none, namely, in between 1:1 and 2:1. As mentioned in the introduction, complex-ratio beat subdivisions in the range of 5:4, 4:3, and 3:2 feature in dance music from Mali (Polak, 2010; Polak & London, 2014). While such ratios prevail in some pieces of repertoire, ratios of about 2:1 prevail in others, and *both* types of rhythm afford equally reliable and precise ensemble synchronization (Polak et al., 2016). Moreover, the more complex ratios are well discriminated and aesthetically preferred by Malian musicians for those pieces that do feature them in performance (Neuhoff et al., 2017). The relativist hypothesis thus would predict that a ratio somewhere around 4:3 ( $\approx 57:43$ ) may constitute a culture-specific prototype for Malian musicians' categorical rhythm perception (see Figure 1, blue triangle in lower tier). Similarly, this hypothesis would also predict a culturally specific category for Bulgarian folk musicians, since much Bulgarian folk music follows metric cycles featuring two beat intervals, long and short, where the ratio of long to short is close to 3:2 (see Djoudjeff, 1931; Goldberg, 2017; Hristov, 1925/1967; Moelants, 2006).

Thus, for the present study we chose participant groups from three countries and four contrasting music-cultural backgrounds: a) Malian drummers specializing in playing for vernacular dance events, b) mostly "classical" percussionists from Germany, whose rhythmic practice tends toward the canonical, simple small-integer ratios, c) Bulgarian musicians who perform folk music, and d) Bulgarian musicians trained in Western classical music.

We chose two tempos to test the hypotheses. As an example of a slow tempo, we adopted the pattern periodicity of 1000 ms, used by Povel (1981) for tapping with periodic two-interval rhythms. We tested three target ratios at that tempo, namely 2:1, 3:2 (an example of a ratio softer than 2:1), and 3:1 (an example of a sharper one); we also included a 1:1 ratio (i.e., isochronous tapping) as a baseline measure. The universalist hypothesis predicts that Povel's results with Westerners should generalize cross-culturally. Specifically, all participants should distort both 3:2 and 3:1 toward 2:1, whereas the reproduction of both 2:1 and 1:1 should involve only comparatively small deviations, if any. The relativist hypothesis, by contrast, predicts that Bulgarian folk musicians should distort the 3:2 ratio to a lesser degree than the other groups do, because 3:2 rhythms play a larger role in Bulgarian folk music than in Western and Malian music at this tempo. The 3:2 rhythm at a periodicity of 1000 ms (600:400 ms) may potentially be heard according to a metric cycle with one long and

one short beat by listeners familiar with Bulgarian folk music repertoires. Comparable meters are very rare in most styles of Western and Malian music.

A fast tempo was considered to compare the perception and production of 2:1 with a particular complex ratio, 58:42. Polak and London (2014) found the 58:42 ratio to be particularly prominent in their study of beat subdivision timings in two different styles of Malian drum ensemble music for vernacular dance events. While the 3:2 (60:40) ratios commonly found in Bulgarian music occur on the beat level, the 58:42 ratio in Malian music occurs on the subdivision level. Beat-level tempos used in the respective Malian genres lie between 80 and 200 bpm. Thus, we chose to use a beat periodicity of 500 ms for the fast tempo. At this periodicity, one can expect participants to perceive the period (500 ms) as an isochronous beat at 120 bpm with two uneven subdivisions.

Note that the 58:42 ratio that we tested is quite close to 4:3 (= 57.14:42.86). The difference between 58:42 and 4:3 ratios would amount to only about four milliseconds per interval at the given periodicity (290:210 ms versus 286:214 ms). While we did not test discrimination, studies on the JND for "swing ratio" differences in two-interval rhythms strongly suggest that this difference would not be discernible (Frane & Shams, 2017; Friberg & Sundström, 2002). For the sake of simplicity, we thus refer to the 58:42 target ratio as "4:3." However, note that this only concerns the ratio of a two-interval rhythm and does *not* at the same time suggest the interpolation of an underlying septuplet subdivision (x . . . x . . .). The latter would be too rapid (approximately 71 ms per element) to suggest a metrically relevant subpulse at the given tempo.

In the context of the fast tempo, the universalist hypothesis predicts that all participant groups (Malian, Bulgarian, and German) will uniformly distort the 4:3 target ratio toward 2:1. By contrast, the relativist hypothesis predicts that only the Bulgarian and German groups will exhibit a distortion toward 2:1, whereas the Malian participants will not.

## Method

### APPARATUS

We played stimuli from an external soundcard attached to a notebook computer and presented the signal to the participants via ear-enclosing studio-monitoring headphones. Participants tapped with one finger on the hard surface of a small idiophonic percussion instrument. They had limited audio feedback from the natural sound of the tapping device, which was dampened by

TABLE 1. *Rhythm Patterns Used as Stimuli*

Tempo	pattern period	L: S ratio	Percentages	L: S quotient	S: L quotient	IOIs (in ms)
Slow	1000 ms	1:1	50:50	1.00	1.00	500:500
		3:2	60:40	1.50	0.67	600:400
		2:1	67:33	2.00	0.50	667:333
		3:1	75:25	3.00	0.33	750:250
Fast	500 ms	1:1	50:50	1.00	1.00	250:250
		≈ 4:3	58:42	1.38	0.72	290:210
		2:1	67:33	2.00	0.50	333:167

TABLE 2. *Participant Groups Used in the Main Experiment*

Country of Residence	Group	Number and gender	Age
Mali	Professional percussionists specializing in folk dance music	$n = 18$ , 1 female	mean = 41, range = 27–55
Germany	Conservatory students of percussion and professional musicians	$n = 14$ , 3 females	mean = 32, range = 20–55
Bulgaria	Conservatory students of folk music	$n = 12$ , 3 females	mean = 23, range = 20–30
Bulgaria	Conservatory students and instructors of Western classical music	$n = 26$ , 13 females	mean = 33, range = 18–67
Total		$N = 70$ , 20 females	mean = 33, range = 18–67

the headphones. We attached a piezoelectric transducer to the interior of the tapping device and recorded both: a) the stimulus from the soundcard, and b) the participant's response from the tapping device via the pickup into an external stereo recorder. The latency and jitter of this apparatus is estimated to be below 1 ms.

#### STIMULI

The stimulus patterns presented in our experiment are summarized in Table 1. All stimuli used a single noisy closed hi-hat sound without definite center frequency, which we manipulated slightly for a particularly sharp tone onset (1 ms from onset to maximum amplitude) and rapid decay. The stimuli do not vary pitch, timbre, loudness, or any dimension other than interval ratio, nor are they primed by metric or other information. Each stimulus sequence starts with the short element and then repeats the long-short rhythmic pattern 38 times.

#### PARTICIPANTS

Table 2 details the participant groups we used. In Mali, we recruited both urban professional and rural semi-professional drummers (players of *jembe* and *dundun*) specializing in percussive ensemble music for local celebratory dance events, such as weddings and staged dance-theatre. These are the musical styles previously studied by Polak and colleagues that motivated our culture-specific hypothesis. In Germany, we tested students of percussion at the Frankfurt University of Music and Performing Arts plus some mature professional musicians who play percussion and other instruments

such as electric bass. Finally, in Bulgaria we distinguished two groups of expert instrumentalists that both consist of conservatory students of music at the Academy of Music, Dance, and Fine Arts in Plovdiv, one group training in Western classical music and the other one in Bulgarian folk music. Members of these groups often play several styles of music and instruments, which include drums and string instruments in the case of the folk music group, and piano, guitar, and clarinet, among several others, in the classical music group. The latter group also included instructors who teach classical music performance at the Academy.

#### PROCEDURE

Participants were instructed by experimenters and local research assistants in Bambara (the *lingua franca* in southern Mali), Bulgarian, or German. They gave their informed consent and received monetary compensation (10–50 EUR, depending on country and level of expertise). After acquainting themselves with the headphones, tapping device, seating position and tapping motion, they tested the audio stimulus in relation to the physical audio feedback of their own tapping response and tried out the task in the example of an isochronous pre-training stimulus. The task was to first listen to the stimulus patterns for a couple of seconds and then start tapping along with each audible sound in the best possible synchrony, continuing as long as the audio stimulus continued to sound.

The experiment was carried out in two parts, one for each of the two tempos. A brief training phase preceded

each part, during which participants first heard short excerpts of each of the stimuli in the set and then carried out practice trials. Participants then performed the set of rhythms three times in three blocks, with a random order of stimuli within blocks. The order of parts (slow versus fast tempo) as well as of stimulus presentation during the training phase within each tempo was counterbalanced across participants. Experimenters cued stimuli on gestural or verbal signs given by the participant, who mostly performed the three to four rhythms per block and the three blocks of a part in a single run, with only a couple of seconds between trials and blocks. The slow tempo part comprised three blocks of four stimuli amounting to twelve trials of about 40 s in length. This typically took 10–15 min to perform. At the fast tempo, three blocks of three stimuli involved nine 20-s trials, which took only four to seven minutes. Instruction, training, and the main experiment at two tempos, a brief interview, and another experiment not reported here were carried out in one session that lasted for about 50–90 min.

#### DATA ANALYSIS

Onsets were extracted from raw audio recordings using a designated Matlab script as follows. We analyzed audio within 8-s windows. For each window, we computed the maximal amplitude of the audio signal and then marked all time points where the signal level crossed a 5% threshold of the maximum, ignoring the second of two onsets separated by less than 120 ms (slow tempo) or 80 ms (fast tempo), because these would clearly result from artifacts, such as an unintentional multiple onset, rather than an intentional tap. We similarly extracted all the stimulus onsets from the recorded audio signal. We paired detected stimulus onsets with response onsets that occurred within a window of 150 ms or 90 ms, respectively, centered around stimulus onsets ( $\pm 75$  or  $\pm 45$  ms tolerance, respectively). Response onsets that did not match these criteria were discarded from further analysis. In general, participants tended to miss no more than three to four (of 38) patterns within a trial, which was mostly due to that in the beginning of each trial, they needed to hear at least one or two iterations of the rhythm before starting to tap.

Our analysis of the response data followed a procedure established in previous tapping studies (e.g., Repp, 2005). We denoted stimulus onsets by  $s_n$ , and response onsets by  $r_n$ . Asynchronies were defined as  $a_n = r_n - s_n$ , where negative asynchrony means that the response onsets precede the stimulus onsets. Inter-response intervals and interstimulus intervals were computed as  $c_n = s_n - s_{n-1}$  and  $d_n = r_n - r_{n-1}$ ,

respectively. The main variable of interest was the percentage of the first (short) interval's duration relative to the duration of the two-element pattern (first plus second interval). We examined separately each pair of consecutively produced response intervals ( $d_{2k-1}, d_{2k}$ ) and calculated the ratio between the first interval within the pair and the summed duration of the two intervals within the same pair. Formally, the relative duration of the first interval ( $o_k$ ) is computed as follows:

$$o_k = 100 \cdot \frac{d_{2k-1}}{d_{2k-1} + d_{2k}},$$

where  $k = 1, \dots, N$  and  $N$  is the number of onset pairs in a trial.

To compute the means of these relative durations as well as the asynchronies reported in the results sections below, we averaged, first, across the repetitions of the two-element pattern in each trial ( $> 30$  repetitions), then across the three trials that participants performed for each stimulus pattern, and finally across the participants per group.

## Results

#### SLOW TEMPO

A  $4 \times 4$  repeated measures ANOVA (ratio  $\times$  group) of the difference of the mean produced relative duration of the first interval from the target relative duration found significant main effects both for ratio,  $F(2.09, 138.28) = 262.21, p < .001, \eta_p^2 = .79$  (Greenhouse-Geisser correction applied here and below) and group,  $F(3, 66) = 4.48, p = .006, \eta_p^2 = .17$ , as well as a significant interaction,  $F(6.28, 138.28) = 3.05, p = .007, \eta_p^2 = .12$ . As shown in Figure 2A, all participant groups distorted the reproduction of both 3:1 and 3:2 targets towards 2:1,  $F(1, 66) = 197.03, p < .001, \eta_p^2 = .75$ ;  $F(1, 66) = 365.62, p < .001, \eta_p^2 = .85$  (Bonferroni correction applied).

Further analyses specific to the different target rhythms showed no significant group effect for the 3:1 rhythmic ratio,  $F(3, 66) = 1.22, p = .90, \eta_p^2 = .05$  (Bonferroni correction applied), nor for the 2:1 ratio,  $F(3, 66) = 5.47, p = .07, \eta_p^2 = .13$ , or 1:1 ratio,  $F(3, 66) = 0.46, p = .71, \eta_p^2 = .02$ , but did indicate a significant group effect for the 3:2 ratio,  $F(3, 66) = 5.94, p = .003, \eta_p^2 = .21$ . The group effect was largely due to the significant difference between the Bulgarian folk musicians and all other groups at the 3:2 rhythmic ratio,  $t(68) = 3.39, p = .002$ .

In addition to the ratio of the two intervals in the response patterns, we also analyzed the variability and

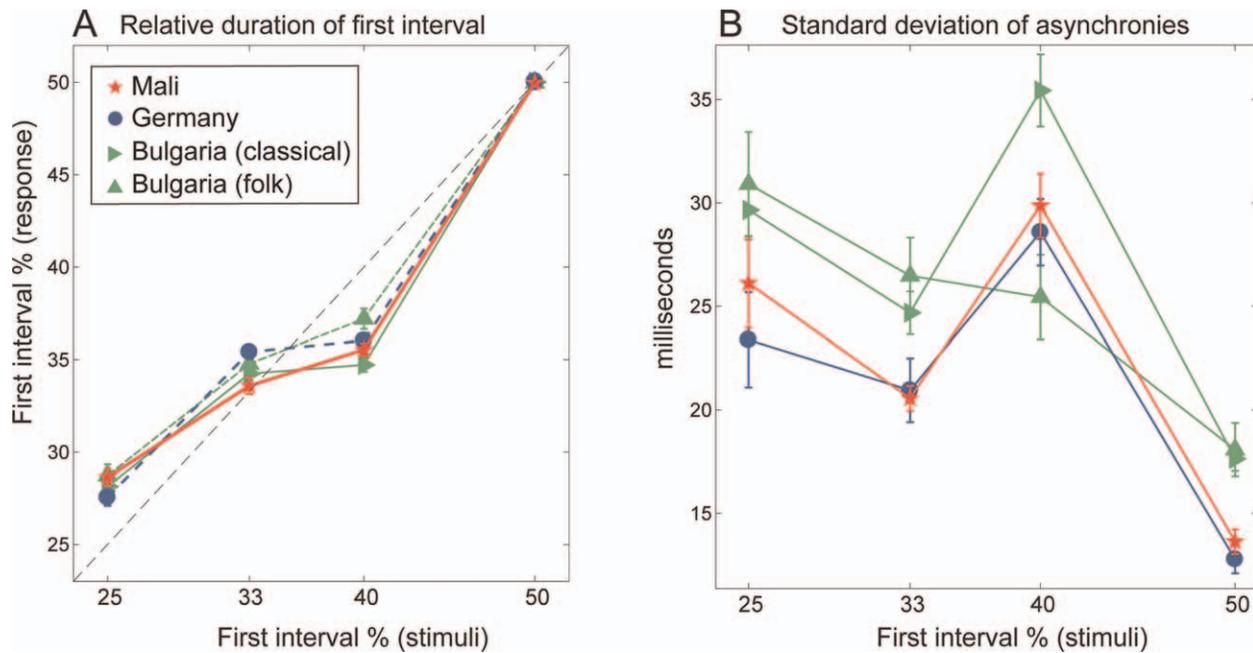


FIGURE 2. A. Responses to the 3:1, 2:1, 3:2, and 1:1 ratios at the slow tempo. The x-axis represents the stimuli (target rhythms) and the y-axis displays the mean values (symbols) and standard errors (error bars) of the corresponding responses. The given values specify the shorter of the two intervals in the rhythmic pattern, in percent of the two-interval pattern duration. Thus, “25” represents 75:25 (3:1), “33” represents 67:33 (2:1), “40” represents 60:40 (3:2), and “50” represents 50:50 (1:1). The dashed diagonal represents identical stimulus and response ratios, i.e., undistorted reproduction. Responses located above that line indicate an increase of the short interval’s proportion, involving a “softening” of the response ratio relative to the target; conversely, responses below the dashed line indicate a decrease of the short interval and thus a “sharpening” of the rhythmic ratio. B. Consistency in the responses to the target ratios at the slow tempo, measured as the standard deviation of asynchronies between corresponding stimulus and response onsets.

magnitude of asynchronies between stimulus and response onsets. The standard deviation of the asynchrony showed main effects of ratio,  $F(2.51, 165.59) = 116.04, p < .001, \eta_p^2 = .64$ , and group,  $F(3, 66) = 6.61, p = .001, \eta_p^2 = .23$ , and a significant interaction,  $F(7.52, 165.59) = 4.24, p < .001, \eta_p^2 = .16$ . The interaction is mainly due to the fact that Bulgarian folk musicians had a significantly smaller standard deviation of the asynchrony (Figure 2B) in the 3:2 ratio compared with Bulgarian classical musicians,  $t(36) = 3.67, p = .001$ , even though their production showed similar standard deviations of the asynchrony at 3:1, 2:1, and 1:1 ratios,  $t(36) = 0.31, p = .76$ ;  $t(36) = 0.67, p = .50$ ;  $t(36) = 0.22, p = .82$ . Mean asynchronies (detailed in Table S-1, see Supplementary Materials section accompanying the online version of this paper), were small in general (between approximately zero and -25 ms) and did not show a significant main effect of group.<sup>2</sup>

<sup>2</sup>The mean asynchrony did not show a significant group effect,  $F(3, 66) = 1.65, p = .19, \eta_p^2 = .07$ , or a group-ratio interaction,  $F(7.98, 173.74) = 1.71, p = .10, \eta_p^2 = .07$ , but we did find a small significant main effect of ratio,  $F(2.63, 173.74) = 8.17, p < .001, \eta_p^2 = .11$ .

## DISCUSSION

The small-integer ratios of 3:1, 2:1, and 3:2 and the pattern periodicity of 1000 ms were chosen because they had been used in classical rhythm tapping studies (e.g., Povel, 1981). Our results successfully replicated Povel’s (1981) main findings across three different countries of residence and four music-cultural groups: Malian drummers, Bulgarian folk musicians, Bulgarian classical musicians, and German percussionists. All groups showed a bias toward the 2:1 ratio, as 3:1 ratios were “softened” in the direction of 2:1 and 3:2 ratios were “sharpened” in the direction of 2:1.

While all groups showed a considerable bias away from the 3:2 ratio toward 2:1, the magnitude of this effect varied among groups. The Bulgarian folk musicians’ bias was smaller in degree than that of their Bulgarian classical, Malian, and German colleagues. In addition, the Bulgarian folk musicians produced this ratio with significantly higher consistency (less variability) than the Bulgarian classical musicians, even though at other target ratios there was no difference between these two groups. Consistent with our relativist hypothesis, we interpret this as an effect of cultural familiarity

with 3:2 rhythms performed in some genres of Bulgarian folk dance music at tempos that include the one used in the stimuli.

The 2:1 ratio itself was not quite exactly reproduced, but rather showed a slight yet significant softening,  $t(69) = -5.02$ ,  $p < .001$ , with a mean for all musicians of around 1.90:1. Such softening of the 2:1 ratio was found by Repp et al. (2011, 2012), but not by other tapping studies (e.g., Povel, 1981). It is worthwhile to report one result of our testing control groups of non-musicians here. In their performance of the 2:1 ratio, nonmusician groups did *not* soften, but tended to slightly sharpen the 2:1 ratio ( $\approx 2.10:1$ ). This difference between the clusters of musician versus nonmusician groups was found to be significant,  $t(126) = -5.41$ ,  $p < .001$  (see the Appendix for further details).

#### FAST TEMPO

At the fast tempo, we used the same method and apparatus as in the slow tempo, but tested a more complex stimulus ratio that is characteristically used in music performance by one of our participant groups, namely, the (approximate) 4:3 ratio common in several styles of Malian percussion music. We also included the 2:1 ratio for comparison.

A  $3 \times 4$  repeated measures ANOVA (ratio  $\times$  group) shows significant effects of ratio,  $F(2, 66) = 169.32$ ,  $p < .001$ ,  $\eta_p^2 = .72$ , and group,  $F(3, 66) = 10.70$ ,  $p < .001$ ,  $\eta_p^2 = .33$ , as well as significant interaction,  $F(3, 66) = 14.97$ ,  $p < .001$ ,  $\eta_p^2 = .72$ . Within the 2:1 ratio there was no significant difference among groups,  $F(3, 66) = 1.06$ ,  $p = .37$ ,  $\eta_p^2 = .46$ . However, as Figure 3A shows, the Malian group did not distort the 4:3 stimulus ratio at all,  $t(17) = -0.36$ ,  $p = .72$ , while all other groups did so to a strong degree,  $t(25) = -11.80$ ,  $p < .001$ ;  $t(11) = -14.98$ ,  $p < .001$ ;  $t(13) = -7.78$ ,  $p < .001$ , for Bulgarian classical, Bulgarian folk, and German percussionists, respectively. A post hoc test showed the Malian group response to be significantly different from all others,  $t(68) = 8.06$ ,  $p < .001$ .

The standard deviation of the asynchrony showed a main effect of ratio,  $F(1.79, 118.25) = 100.86$ ,  $p < .001$ ,  $\eta_p^2 = .60$ , of group,  $F(3, 66) = 12.49$ ,  $p = .001$ ,  $\eta_p^2 = .36$ , and a significant interaction,  $F(5.37, 118.25) = 12.39$ ,  $p < .001$ ,  $\eta_p^2 = .36$ . The mean asynchrony showed significant ratio,  $F(1.72, 113.94) = 20.00$ ,  $p < .001$ ,  $\eta_p^2 = .23$ , and group effects,  $F(3, 66) = 1.65$ ,  $p = .19$ ,  $\eta_p^2 = .07$ , as well as an interaction,  $F(5.17, 113.94) = 8.25$ ,  $p < .001$ ,  $\eta_p^2 = .27$ . The details of mean asynchronies are presented in Table S-1 (Supplementary Materials section online).

The Malian drummers' quite precise reproduction of the 4:3 target ratio stands in stark contrast with all other

groups' strongly biased response behavior. This difference is confirmed by the fact that the Malian group's variability, measured as the standard deviation of asynchronies between stimulus and response, is considerably smaller than the other groups' variability in tapping to the 4:3 target,  $t(68) = -6.08$ ,  $p < .001$ .

The Malian group's unbiased reproduction of the ratio as well as the low variability in this behavior are consistent with the culture-relativist hypothesis of a distinct perceptual prototype engendered by the Malians' particular musical enculturation. However, there are several alternative explanations of our findings. First, one may suspect that the Malian group outperforms other groups in the response to the 4:3 target ratio because of a generally superior capacity for sensorimotor synchronization. However, this could not explain why the Malian drummers then did not outperform the other musicians at other ratios in the fast tempo and at the slow tempo, too. For instance, in the 1:1 tapping task at the fast tempo, Malians displayed standard deviations from the target not significantly different from the German group,  $t(30) = 1.38$ ,  $p = .17$ .

Furthermore, we can rule out the possibility that the Malian musicians' group mean is a result of highly variable responses that incidentally average to an apparently unbiased result. Indeed, the standard deviation of asynchronies between stimulus and response when tapping to the 4:3 target is smaller for the Malian compared with the other groups,  $t(68) = -6.08$ ,  $p < .001$ . Moreover, Figure 3B shows that the Bulgarian and German groups' responses are considerably less consistent (more variable) when tapping to the 4:3 stimulus than to the 2:1 target,  $t(36) = 1.33$ ,  $p = .19$ . By contrast, the Malian group's degree of consistency is roughly comparable across both ratios. To further verify that our results are not somehow an artifact of the averaging procedure, we plotted the histogram of all the first intervals within each of the response two-interval pairs in the experiment (Figure 3C). We emphasize that these histograms were obtained without any averaging as all single raw data repetitions were used directly. The histogram indicates that the difference of the Malian percussionists from all other groups is not a product of averaging a massively wider distribution. Rather, the Malian groups' distribution in its entirety is shifted to the right.

A last alternative explanation to be considered is that our results may concern an elongation or compression, respectively, or an increased inaccuracy of one of the two intervals rather than the ratio (relative duration) of the two elements in the response pairs. Yet neither the mean nor the variability of the asynchrony was significantly different between the first and second interval, as

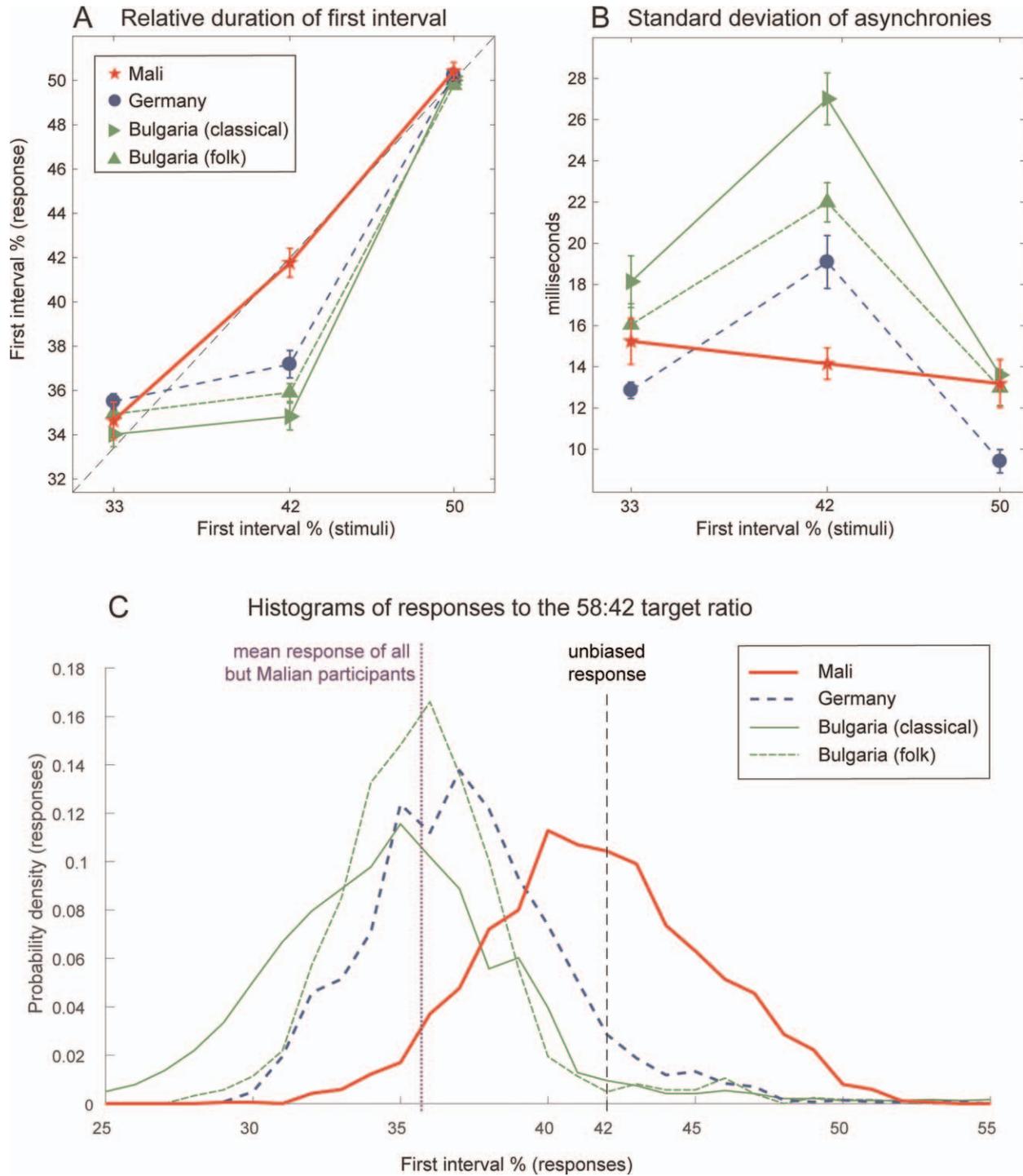


FIGURE 3. A. Responses to the 2:1, 4:3, and 1:1 ratios at the fast tempo. The x-axis represents the stimuli (target rhythms) and the y-axis displays the mean values (symbols) and standard errors (error bars) of the corresponding responses. The given values specify the shorter of the two intervals in the rhythmic pattern, in percent of the two-interval pattern duration. The dashed diagonal represents identical stimulus and response ratios, i.e., undistorted reproduction. B. Consistency in the responses to the target ratios at the fast tempo, measured as the standard deviation of asynchronies between corresponding stimulus and response onsets. C. Histogram of responses to the 4:3 target ratio at the fast tempo. The black dashed vertical line shows the unbiased response, and the purple dotted vertical line shows the mean response of the aggregated Bulgarian and German groups, i.e., the cluster of all groups except the Malian one.

indicated by the lack of a significant main effect of interval within a group  $\times$  ratio  $\times$  interval 3-way ANOVA; mean:  $F(1, 66) = 0.79, p = .38, \eta_p^2 = .01$ ; std:  $F(1, 66) = 0.66, p = .41, \eta_p^2 = .01$  (see Table S-2 in the Supplementary Materials section online for details on the mean values and standard deviations of the asynchronies separated for the first and second onsets, respectively). This is consistent with the idea that the response behavior was driven by the two-interval pattern ratio holistically rather than by the duration or inaccuracy of one of the two intervals alone.

In sum, various additional analyses falsify alternative explanations of the Malian groups' precise and consistent reproduction of the 4:3 target stimulus, and are consistent with the hypothesis of a culture-specific perceptual prototype.

#### DISCUSSION

We tested the hypothesis that a culture-specific perceptual category would be evident when musician participants were given a rhythm reproduction task that specifically engages that category. We found support for this hypothesis using a two-interval rhythm whose elements relate by approximately 4:3 at a tempo of 120 bpm, a rhythmic pattern that is characteristic of various repertoires and styles of music for social dance in Mali. The German and Bulgarian musicians' distortion of the 4:3 ratio towards 2:1 replicated the same groups' response to the similar 3:2 ratio in the slower tempo. By contrast, Malian musicians did not distort the 58:42 ratio towards 2:1, but reproduced it with a high degree of accuracy and consistency.

#### General Discussion

We tested musician groups from diverse cultural backgrounds to see if there were systematic differences in their reproductions of two-interval rhythms at two tempos. We found such differences, supporting the relativist hypothesis that categorical rhythm perception depends on culture. This hypothesis assumed that when presented with a stimulus relevant to a culturally specific rhythmic performance prototype, musicians whose cultural background included that performance prototype would show advantageous effects of using a corresponding cognitive prototype in their synchronization behavior, while musicians whose cultural background did not include that prototype would not have such a corresponding cognitive prototype at their disposal. Consistent with this assumption, we found Bulgarians trained in folk music performance to differ in their synchronization with the 3:2 rhythm, when presented at a slow

tempo, from both their colleagues in Bulgaria with Western classical training and from musicians in Germany and Mali. The Bulgarian folk musicians displayed a weaker bias away from the 3:2 rhythm and showed less variability in their behavior than did all other groups. This is in line with several studies of participants with cultural backgrounds in the Balkans, Turkey, and India that evidence positive influence of cultural familiarity with 3:2-based rhythms on processing of such rhythms (Hannon & Trehub, 2005a; Hannon et al., 2012; Kalendar et al., 2013; Ullal-Gupta et al., 2014).

In addition, we tested the potential influence of culture-dependent constraints at a fast tempo by presenting participants with a relatively complex two-element rhythm (approximately 4:3) that was highly familiar to one group (Malian participants) but not to others. Consistent with the relativist hypothesis, we found that the Malians performed the 4:3 rhythm task with a high degree of fidelity to the target ratio, whereas all other groups distorted the 4:3 stimulus pattern in the direction of 2:1. This suggests that the Malian musicians—but not the Bulgarian and German musicians—have a cognitive category whose prototype lies in the vicinity of 4:3.

Along with this cultural variation, we also found a degree of cross-cultural similarity in the rhythmic behavior of the groups we studied. First, the 1:1 ratio (isochrony) was performed with great fidelity by all groups at both tempos. Secondly, a slight bias (softening) in the reproduction of the 2:1 ratio occurred identically across cultural groups. Contemporary conceptualizations of “statistical universals” in music argue that musical features vary in the frequency of occurrence within and across cultures (Brown & Jordania, 2013; Savage et al., 2015). From this perspective, our findings are consistent with the assumption that rhythmic prototypes tied to the simplest small-integer ratios may be a statistical universal of high frequency.

#### THE POSSIBLE STATUS OF THE 4:3 PROTOTYPE

The recent Bayesian conceptualization of perceptual categories based on perceptual priors allows for continuity and “fuzzy” overlap between perceptual categories (Feldman et al., 2009). According to this concept, overlapping categories occur when the a priori probability of intermediate percepts is not zero. For instance, Jacoby and McDermott (2017) demonstrated that the categories characterized by the prototypes that lie near 2:1 and 3:1 overlap and interact with each other rather than being completely distinct. Analogously, it is quite likely that the closely neighboring 4:3 and 2:1 prototypes we found in our study are not discretely separate from each other. Rather, one may assume that the categories tied to the

4:3 and 2:1 prototypes in our Malian participants' case overlap in a way roughly comparable to the way the 2:1 and 3:1 categories overlap in Jacoby and McDermott's (2017) study (see the histogram in our Figure 1).

Our data are consistent with at least two potential descriptions of the alleged relationship between the 2:1 and 4:3 categories. One possibility is that they are overlapping yet independent. Malian participants exposed to a musical repertoire and style that frequently displays rhythms with ratios close to 4:3 may have developed over time a corresponding perceptual category. This category would define a segment of the rhythmic space that otherwise, as in Jacoby and McDermott's study of Western participants, is left ambiguous (see the valley between the 1:1 and 2:1 prototypes in the histogram in our Figure 1). An alternative possibility is that 4:3 and 2:1 are two related "flavors" of the same category (long-short), that is, subcategories. Perceptual categories can nest into each other in hierarchical ways. For example, turquoise and olive may be seen as distinct colors and as different shades of green at the same time. When compared one against the other, most English speakers would not find it difficult to recognize them correctly as distinct categories, though both make part of a more general category, namely, green, which is more robust across individuals and cultures (Berlin & Kay, 1969). Notably, each of the pieces of repertoire of Malian music motivating our research hypothesis shows a stable relation with *either* the 2:1 *or* the 4:3 category, as particular pieces in the respective repertoires employ particular subdivision ratios (Polak, 2010; Polak & London, 2014; Polak et al., 2016). In other words, in the performance practice in question the 2:1 and 4:3 prototypes do not co-occur within a piece, but rather appear in separate musical situations. One may thus assume that the 4:3 and 2:1 prototypes, though discriminable from each other in the tested laboratory setting, ecologically are instead used as subcategories of a single primary category (long-short). Further experimentation will be needed to test these hypothetical explanations. For example, future research may test the degree of overlap between the 4:3 and 2:1 prototypes by characterizing the perceptual space between and around them using fine-grained intermediate stimuli and iterative reproduction.

#### RATIO AND CONTEXT

A key factor in the Malians' performance is tempo. We presented similar rhythms at two tempos and the Malian 4:3 prototype occurred only at the more rapid tempo, which is a key characteristic of the Malian performance practice. This suggests that the Malian 4:3 prototype is part of a "tempo-metrical type" (London,

2012); that is, a particular template for meter that occurs only within a specific range of tempos. Relations among metrical levels (beat versus subdivision) and rhythmic elements (upbeat versus downbeat) can change substantially as tempo is increased or decreased (London, Humberg, & Cross, 2009). The Malian 4:3 performance pattern is manifest on the level of subdivision of an otherwise isochronous beat (Polak, 2010; Polak & London, 2014); in our view, this is why a stimulus that features this ratio needs to be sufficiently fast to activate the corresponding cognitive prototype. At a slow tempo, the rhythmic ratios will tend to engage mechanisms of beat induction and interpolation (Desain & Honing, 1999; Honing, 2013; London, 2012; Parncutt, 1994; Repp, 2003, 2008; Tal et al., 2017) rather than to elicit the feeling of a beat subdivision. We propose it is this strong effect of tempo on meter and rhythm perception that explains the stark contrast in our Malian participants' rhythmic behavior between slow and fast tempos. The hypothesized effect of a culture-specific 3:2-prototype in the Bulgarian folk musicians, too, emerged only at one of two tested tempos; namely, at a tempo that falls within the range of typical performance tempos of the hypothesized prototype.

We speculate that culture-specific categories for rhythm perception beyond those based on the very simplest integer ratios may generally depend on combinations of ratio (durational proportions) plus other factors relevant for the corresponding music-cultural performance contexts. This may be tempo, as suggested by the present study, yet other contextual factors may be relevant, too, such as sound, melody, motional gesture, etc. If this is true, the context-poor stimuli typically used in rhythm reproduction experiments might not suffice to realistically study categorical rhythm perception. Future research may explore the potential value of more ecologically valid, context-rich stimuli. For instance, the target ratios for synchronization tasks may be embedded in melodic excerpts instead of equitone series.

The potential problem of context-poor stimuli for activating effects of experience may relate to a certain contradiction in the field of categorical rhythm perception. On the one hand, the relevance of the simplest integer ratios is strongly emphasized; on the other hand, these alone can hardly explain the complex rhythmic structures that feature in many musical practices. This was noted previously by Dirk-Jan Povel, despite the fact that his research supported Paul Fraisse's influential proposition that there are only two basic categories for rhythm perception with prototypes at 1:1 and 2:1: "The suggestion that the perception and production of rhythms could be understood by an internal

representation that allows only two distinct durations seems too simple” (Povel, 1981, p. 3).

#### CULTURE AND EXPERTISE

Our study focused on testing musical experts, as we needed to have participants with: a) the high sensorimotor skills that are required to tap at fast tempos, and b) a high degree of enculturation with the particular rhythmic ratios that were of interest. Using musicians allowed us to disaggregate the contributions of general expertise versus culturally specific expertise: All musicians tested were highly trained and demonstrated strong skills in basic sensorimotor synchronization. However, the two cases of intergroup variation in response behavior (Bulgarian folk musicians at 3:2 at the slow tempo, Malian drummers at 4:3 at the fast tempo) both depended on a high level of musical experience *and* experience in a *particular* cultural and stylistic milieu.

We made our operational definition of “musician” based upon participant self-identification, current and ongoing public performance activities, and with a requirement that they were instrumentalists. This also implies an operational definition of “nonmusician,” i.e., someone who fails to meet all of the criteria. To assess the construct validity of our definitions, we administered the tapping experiment at the slow tempo to “nonmusicians” from Mali, Bulgaria, and Germany and found significant differences between musician and nonmusician groups (see Figures A1 and A2 in the Appendix). Note that our definition does not preclude nonmusicians from having high degrees of musicality (Bigand & Poulin-Charronnat, 2006) and familiarity with a particular musical style; some of our Malian and all of our Bulgarian nonmusician participants were dancers with indeed a high degree of familiarity and listening and interaction expertise with their respective musical styles. Nonetheless, they did not perform as well on the tapping task at the slow tempo, and according to our pilot experiment often found it difficult to perform the task at the fast tempo at all. Previous research relying on Western participants has shown that experience in playing instruments enhances performance in sensorimotor synchronization tasks (Manning & Schutz, 2016; Matthews, Thibodeau, Gunther, & Penhune, 2016; Repp, 2010). The present study complements this research by evidencing its cross-cultural validity. More generally, this gives hope that distinguishing musician and nonmusician groups can make sense in at least some cross-cultural research paradigms and settings; however, the problem of cross-cultural equivalence of concepts may need to be addressed in each new experimental context and sampling of cases.

#### CULTURE AND PERCEPTUAL LEARNING

A consensus of previous literature has assumed that production biases in tapping experiments result largely from categorical perception (see Introduction). However, the method of rhythm reproduction through sensorimotor synchronization does not allow us to exclude that production (e.g., motor) constraints contributed to our results. In addition, it is possible that different groups of participants would show a variable degree of reliance on production while tapping. More cross-cultural research directly addressing the relation of perception and production will be required to resolve the relative contribution of these factors. It appears safe to assume, though, that perception at least contributes to the observed behavior. Thus, the finding of a culture-specific 4:3 prototype for two-element rhythms arguably suggests a case of large-scale perceptual category differentiation (Goldstone, 1998; Goldstone & Hendrickson, 2010) at the level of culture.

From an ecological perspective, our findings of culture-specific perceptual prototypes do not come as a surprise. Correlations between culture-specific perception and performance traditions (genres, repertoires, and styles) seem plausible, since humans develop perceptual skills in the process of exploring and learning to perceive their environments (Gibson, 1963, 1969; Gibson & Gibson, 1955; Goldstone, Landy, & Son, 2010; Goldstone et al., 2015), and parts of these environments are human-made. The latter point has recently been emphasized in connection with the biological concept of human cultural niche construction, according to which humans have *not* simply inhabited and adapted to stable environments since prehistoric times; rather, humans and their environments tend to dynamically shape each other in feedback loops at various time-scales, from ontogenetic to historical to evolutionary (Laland & O’Brien, 2011). Psychologists, archeologists, anthropologists, and philosophers, among others, have proposed a strong role for cultural niche construction in human development, culture change, cognition, and arts (Bertolotti & Magnani, 2017; Flynn, Laland, Kendal, & Kendal, 2013; Gauvain, 2000; Heft, 2007; Kendal, 2011; Menary, 2015; Portera, 2016; Riede, 2011; Schultz, 2014; 2016; Stotz, 2010). Along these lines of research, conceiving of music as a human-made aspect of the human environment suggests the possibility that perception and performance structures of music may culturally co-evolve. Future research thus may consider the notion of *perceptual niche construction* as a possible explanation for cases of cultural variation or change in music perception.

Our evidence for a 4:3 rhythmic prototype is inconsistent with the assumption that the wide distribution of the 1:1 and 2:1 prototypes results from biological constraints that preclude rhythmic categories characterized by more complex ratios. We thus second Norenzayan and Heine (2005) in their cautioning against an often implicit conflating of universals with innateness or genetic determination. This is not to deny that the simplicity of the smallest integer ratios is of advantage to rhythm perception and cognition, nor is it to deny the relevance of their wide distribution. Yet while innate disposition certainly is a plausible candidate for the explanation of certain universals, they may also result from similar problems suggesting similar solutions, which can both develop independently and spread in the cultural frameworks of social practice and learning. It will thus require more research to understand the potentially universal spread of the 1:1 and 2:1 prototypes, and the strong role of small-integer ratios for rhythm perception in general.

### Interdisciplinary Implications

The study of music perception and cognition is increasingly interested in the extent of cross-cultural variation that may or may not accompany the well-documented variability of music performance practices (Stevens, 2012; Stevens & Byron, 2016). Cross-cultural research perspectives are a particularly welcome means of working against systematic biases inherent in the usage of non-representative participant populations (cf. Henrich, Heine, & Norenzayan, 2010a, 2010b). Beller, Bender, and Medin (2012) recently argued that cognitive science does *not* need socio-cultural anthropology as a team member in this endeavor to become more culturally sensitive. The ethnographic case studies usually produced by anthropologists and ethnomusicologists tend to be dismissed by scientists as inaccessible and unreliable. Yet their qualitative knowledge about culture, cultures, and cross-cultural variation can inspire experimental research with interesting hypotheses, guide the sampling of cases, provide access to participant populations, help to choose and modify experimental designs, and interpret responses (cf. Astuti & Bloch, 2010; Baumard & Sperber, 2010; Shweder, 2010; Whitehouse & Cohen, 2012). The present study is a case in point. Without specific knowledge of Malian musical culture and practice developed in a series of anthropological, ethnomusicological, music-analytical, and empirical-musicological case studies, we would not have recognized the presence of a culturally specific rhythmic prototype at the metric level of beat subdivision in West Africa.

Consequently, we would not have developed the hypothesis that this prototype may depend on the combination of a particular ratio *and* a specific tempo range. Now imagine what would have happened if a team of culturally naïve scholars had simply tried to replicate in Mali an established paradigm from the literature—as we did in the slow tempo part of our experiment. Such a team would have found a striking degree of coherence overall and would probably have argued that this validates the universalist hypothesis and invalidates the relevance of the culturally relativist perspective. Our interdisciplinary approach demonstrates the pitfalls of such naïveté. Identifying hypothetically culture-specific variations in perceptual systems requires detailed knowledge of the specific musical environments to be studied. It will often require a sensitive adjustment of experimental parameters to evoke effects of cultural familiarity that could easily be overlooked in the context of the ever-pressing scientific striving to minimize variables for the sake of control. Thus, however large the methodological and epistemological differences between the sciences and the humanities may be, understanding the role of culture in human perception and cognition will be hard to achieve without their collaboration.

### Author Note

Authors Polak and Jacoby contributed equally to the research. Polak led the study and Jacoby analyzed the data. Experiments were conducted by Polak, Goldberg, Fischinger, and Holzapfel. Polak and Jacoby drafted the paper, which was edited by London, Goldberg, Fischinger, and Holzapfel.

We would like to thank Noumouké Doumbia, Dr. Salabary Doumbia, and Madu Jakite (Bamako), Dr. Ozan Baysal (Istanbul), Elena Felker (Frankfurt), and Dr. Lyuben Dossev (Plovdiv) for their assistance in conducting experiments.

The *German Research Foundation* (DFG, research grant PO-627/6-3, grant-holder Polak) and the *Cologne University for Music and Dance* (HfMT Köln) generously funded the coordination of the study and data acquisition in Mali, Turkey and Bulgaria; data acquisition in Germany and writing the paper was funded by the *Max Planck Institute for Empirical Aesthetics* (MPIEA). Author Polak changed from HfMT Köln to MPIEA during the research.

*Correspondence concerning this article should be addressed to* Rainer Polak, Music Department, Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt am Main, Germany. E-mail: rainer.polak@ae.mpg.de

## References

- AMSELLE, J.-L., & M'BOKOLO, E. (2009). *Au coeur de l'ethnie: Ethnies, tribalisme et état en Afrique* [At the heart of ethnicity: Ethnic groups, tribalism and the state in Africa] (2nd ed.). Paris, France: La Découverte.
- ASTUTI, R., & BLOCH, M. (2010). Why a theory of human nature cannot be based on the distinction between universality and variability: Lessons from anthropology. *Behavioral and Brain Sciences*, 33, 83–84. DOI: 10.1017/S0140525X10000026
- BARTH, F. (1998). *Ethnic groups and boundaries: The social organization of culture difference*. Long Grove, IL: Waveland Press.
- BASHKOW, I. (2004). A neo-Boasian conception of cultural boundaries. *American Anthropologist*, 106, 443–458. DOI: 10.1525/aa.2004.106.3.443
- BATES, E. (2011). *Music in Turkey: Experiencing music, expressing culture*. New York: Oxford University Press.
- BAUMARD, N., & SPERBER, D. (2010). Weird people, yes, but also weird experiments. *Behavioral and Brain Sciences*, 33, 84–85. DOI: 10.1017/S0140525X10000038
- BELLER, S., BENDER, A., & MEDIN, D. L. (2012). Should anthropology be part of cognitive science? *Topics in Cognitive Science*, 4, 342–353. DOI: 10.1111/j.1756-8765.2012.01196.x
- BENADON, F. (2006). Slicing the beat: Jazz eighth-notes as expressive microrhythm. *Ethnomusicology*, 50, 73–98.
- BERLIN, B., & KAY, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley, CA: University of California Press.
- BERTOLOTI, T., & MAGNANI, L. (2017). Theoretical considerations on cognitive niche construction. *Synthese*, 194, 4757–4779. DOI: 10.1007/s11229-016-1165-2
- BIGAND, E., & POULIN-CHARRONNAT, B. (2006). Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. *Cognition*, 100, 100–130. DOI: 10.1016/j.cognition.2005.11.007
- BOURDIEU, P. (1984). *Distinction: A social critique of the judgement of taste*. Cambridge, MA: Harvard University Press.
- BOYD, R., & RICHERSON, P. J. (2005). *The origin and evolution of cultures*. Evolution and cognition. Oxford, UK: Oxford University Press.
- BRADLOW, A. R., AKAHANE-YAMADA, R., PISONI, D. B., & TOHKURA, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception and Psychophysics*, 61, 977–985. DOI: 10.3758/BF03206911
- BRĂILOIU, C. (1984). Aksak rhythm. In A. L. Lloyd (Ed.), *Problems of ethnomusicology* (pp. 133–167). Cambridge, UK: Cambridge University Press. (Reprinted from *Revue de Musicologie*, 33, 71–108, 1951)
- BROWN, S., & JORDANIA, J. (2013). Universals in the world's musics. *Psychology of Music*, 41, 229–248. DOI: 10.1177/0305735611425896
- BRUMANN, C. (1999). Writing for culture: Why a successful concept should not be discarded. *Current Anthropology*, 40(S1), S1–S27. DOI: 10.1086/200058
- BUTTERFIELD, M. W. (2011). Why do jazz musicians swing their eighth notes? *Music Theory Spectrum*, 33, 3–26. DOI: 10.1525/mts.2011.33.1.3
- CAMERON, D. J., BENTLEY, J., & GRAHN, J. A. (2015). Cross-cultural influences on rhythm processing: Reproduction, discrimination, and beat tapping. *Frontiers in Psychology*, 6, 366. DOI: 10.3389/fpsyg.2015.00366
- CLARKE, E. F. (1985). Structure and expression in rhythmic performance. In P. Howell, R. West, & I. Cross (Eds.), *Musical structure and cognition* (pp. 209–236). London, UK: Academic Press.
- CLARKE, E. F. (1987). Categorical rhythm perception: An ecological perspective. In A. Gabrielsson (Ed.), *Action and perception in rhythm and music* (pp. 19–33). Stockholm, Sweden: Royal Swedish Academy of Music.
- CLER, J. (1994). Pour une théorie de l'aksak [Towards a theory of aksak]. *Revue de Musicologie*, 80, 181–210. DOI: 10.2307/947054
- DESAIN, P., & HONING, H. (1999). Computational models of beat induction: The rule-based approach. *Journal of New Music Research*, 28, 29–42. DOI: 10.1076/jnmr.28.1.29.3123
- DESAIN, P., & HONING, H. (2003). The formation of rhythmic categories and metric priming. *Perception*, 32, 341–365. DOI: 10.1068/P3370
- DJOUJJEFF, S. (1931). *Rythme et mesure dans la musique populaire bulgare* [Rhythm and meter in Bulgarian folk music]. Paris, France: Librairie Ancienne Champion.
- DRAKE, C. (1993). Reproduction of musical rhythms by children, adult musicians, and adult nonmusicians. *Perception and Psychophysics*, 53, 25–33. DOI: 10.3758/BF03211712
- DRAKE, C., & BERTRAND, D. (2001). The quest for universals in temporal processing in music. *Annals of the New York Academy of Sciences*, 930, 17–27. DOI: 10.1111/j.1749-6632.2001.tb05722.x
- DRAKE, C., & EL HENI, J. B. (2003). Synchronizing with music: Intercultural differences. *Annals of the New York Academy of Sciences*, 999, 429–437. DOI: 10.1196/annals.1284.053
- DURING, J. (1997). Rythmes ovoïdes et quadrature du cycle [Ovoid rhythms and the squaring of the circle]. *Cahiers de musiques traditionnelles*, 10, 17–36.
- ELFENBEIN, H. A., & AMBADY, N. (2003). Cultural similarity's consequences: A distance perspective on cross-cultural differences in emotion recognition. *Journal of Cross-Cultural Psychology*, 34, 92–110. DOI: 10.1177/0022022102239157
- EPSTEIN, D. (1985). Tempo relations: A cross-cultural study. *Music Theory Spectrum*, 7, 34–71. DOI: 10.2307/745880

- ESSENS, P., & POVEL, D.-J. (1985). Metrical and non-metrical representations of temporal patterns. *Perception and Psychophysics*, 37, 1–7.
- FELDMAN, N. H., GRIFFITHS, T. L., & MORGAN, J. L. (2009). The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological Review*, 116, 752–782. DOI: 10.1037/a0017196
- FLYNN, E. G., LALAND, K. N., KENDAL, R. L., & KENDAL, J. R. (2013). Target article with commentaries: Developmental niche construction. *Developmental Science*, 16, 296–313. DOI: 10.1111/desc.12030
- FRAISSE, P. (1956). *Les structures rythmiques*. Louvain, Belgium: Publications Universitaires de Louvain.
- FRAISSE, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (1st ed., pp. 149–180). New York: Academic Press.
- FRANE, A. V., & SHAMS, L. (2017). Effects of tempo, swing density, and listener's drumming experience, on swing detection thresholds for drum rhythms. *Journal of the Acoustical Society of America*, 141, 4200–4208. DOI: 10.1121/1.4984285
- FRIBERG, A., & SUNDSTRÖM, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception*, 19, 333–349. DOI: 10.1525/mp.2002.19.3.333
- GABRIELSSON, A., BENGTSSON, I., & GABRIELSSON, B. (1983). Performance of musical rhythm in 3/4 and 6/8 meter. *Scandinavian Journal of Psychology*, 24, 193–213. DOI: 10.1111/j.1467-9450.1983.tb00491.x
- GAUVAIN, M. (2000). Niche construction, social co-construction, and the development of the human mind. *Behavioral and Brain Sciences*, 23, 153. DOI: 10.1017/S0140525X00302412
- GERISCHER, C. (2003). *O Suingue Baiano: Mikrorhythmische Phänomene in baianischer Perkussion* [The Bahian swing: Microrhythmic phenomena in Bahian percussion]. Frankfurt, Germany: Peter Lang.
- GERISCHER, C. (2006). *O suingue baiano: Rhythmic feeling and microrhythmic phenomena in Brazilian percussion*. *Ethnomusicology*, 50, 99–119.
- GIBSON, E. J. (1963). Perceptual learning. *Annual Review of Psychology*, 14, 29–56. DOI: 10.1146/annurev.ps.14.020163.000333
- GIBSON, E. J. (1969). *Principles of perceptual learning and development*. East Norwalk, CT: Appleton-Century-Crofts.
- GIBSON, J. J., & GIBSON, E. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review*, 62, 32–41. DOI: 10.1037/h0048826
- GOLDBERG, D. (2015). Timing variations in two Balkan percussion performances. *Empirical Musicology Review*, 10, 305–328. DOI: 10.18061/emr.v10i4.4884
- GOLDBERG, D. (2017). *Bulgarian meter in performance* (Unpublished doctoral dissertation). Yale University, New Haven, CT.
- GOLDSTONE, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, 49, 585–612. DOI: 10.1146/annurev.psych.49.1.585
- GOLDSTONE, R. L., & HENDRICKSON, A. T. (2010). Categorical perception. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1, 69–78. DOI: 10.1002/wcs.26
- GOLDSTONE, R. L., LANDY, D. H., & SON, J. Y. (2010). The education of perception. *Topics in Cognitive Science*, 2, 265–284. DOI: 10.1111/j.1756-8765.2009.01055.x
- GOLDSTONE, R. L., LEEUW, J. R. DE, & LANDY, D. H. (2015). Fitting perception in and to cognition. *Cognition*, 135, 24–29. DOI: 10.1016/j.cognition.2014.11.027
- GOODY, J. (1992). Culture and its boundaries: A European view. *Social Anthropology*, 1, 9–32. DOI: 10.1111/j.1469-8676.1992.tb00238.x
- GUPTA, A., & FERGUSON, J. (1992). Beyond “culture”: Space, identity, and the politics of difference. *Cultural Anthropology*, 7, 6–23. DOI: 10.1525/can.1992.7.1.02a00020
- HANNON, E. E., & TREHUB, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science*, 16, 48–55. DOI: 10.1111/j.0956-7976.2005.00779.x
- HANNON, E. E., & TREHUB, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences*, 102, 12639–12643. DOI: 10.1073/pnas.0504254102
- HANNON, E. E., SOLEY, G., & ULLAL-GUPTA, S. (2012). Familiarity overrides complexity in rhythm perception: A cross-cultural comparison of American and Turkish listeners. *Journal of Experimental Psychology: Human Perception and Performance*, 38, 543–548. DOI: 10.1037/a0027225
- HANNON, E. E., VANDEN BOSCH DER NEDERLANDEN, C. M., & TICHKO, P. (2012). Effects of perceptual experience on children's and adults' perception of unfamiliar rhythms. *Annals of the New York Academy of Sciences*, 1252, 92–99. DOI: 10.1111/j.1749-6632.2012.06466.x
- HARNAD, S. R. (1990). *Categorical perception: The groundwork of cognition*. Cambridge, UK: Cambridge University Press.
- HAUGEN, M. R. (2016). Investigating periodic body motions as a tacit reference structure in Norwegian telespringar performance. *Empirical Musicology Review*, 11, 272–294. DOI: 10.18061/emr.v11i3-4.5029
- HEFT, H. (2007). The social constitution of perceiver-environment reciprocity. *Ecological Psychology*, 19, 85–105.
- HENRICH, J., HEINE, S. J., & NORENZAYAN, A. (2010a). The weirdest people in the world? *Behavioral and Brain Sciences*, 33, 61–83. DOI: 10.1017/S0140525X0999152X
- HENRICH, J., HEINE, S. J., & NORENZAYAN, A. (2010b). Most people are not WEIRD. *Nature*, 466, 29. DOI: 10.1038/466029a
- HENRICH, J., & TENNIE, C. (2017). Cultural evolution in chimpanzees and humans. In M. N. Muller, R. W. Wrangham, & D. R. Pilbeam (Eds.), *Chimpanzees and human evolution*. Cambridge, MA: Belknap Press.

- HOLZAPFEL, A. (2015). Relation between surface rhythm and rhythmic modes in Turkish makam music. *Journal of New Music Research*, 44, 25–38. DOI: 10.1080/09298215.2014.939661
- HONING, H. (2013). Structure and interpretation of rhythm in music. In D. Deutsch (Ed.), *The psychology of music* (3rd ed., pp. 369–404). London, UK: Academic Press. DOI: 10.1016/B978-0-12-381460-9.00009-2
- HRISTOV, D. (1967). Metrichnite i ritmichnite osnovi na bŭlgarskata narodna muzika. In V. Krŭstev (Ed.), *Muzikalno-teoretichsko i publitsistichsko nasledstvo* [Music-theoretical and publicistic legacy] (Vol. 1, pp. 33–98). Sofia: Bŭlgarskata Akademiya na Naukite. (Original work published 1925)
- IVERSEN, J. R., PATEL, A. D., & OHGUSHI, K. (2008). Perception of rhythmic grouping depends on auditory experience. *Journal of the Acoustical Society of America*, 124, 2263–2271. DOI: 10.1121/1.2973189
- JACOBY, N., & McDERMOTT, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Current Biology*, 27, 359–370. DOI: 10.1016/j.cub.2016.12.031
- JANKOWSKY, R. C. (2013). Rhythmic elasticity, metric ambiguity, and ritual teleology in Tunisian stambeli. *Analytical Approaches to World Music*, 3(1), 34–61. Retrieved from [http://www.aawmjournal.com/articles/2014a/Jankowsky\\_AAWM\\_Vol\\_3\\_1.pdf](http://www.aawmjournal.com/articles/2014a/Jankowsky_AAWM_Vol_3_1.pdf)
- JOHANSSON, M. (2009). *Rhythm into style: Studying asymmetrical grooves in Norwegian folk music* (Unpublished doctoral thesis). University of Oslo, Oslo, Norway.
- JOHANSSON, M. (2017). Non-isochronous musical meters: Towards a multidimensional model. *Ethnomusicology*, 61, 31–51. DOI: 10.5406/ethnomusicology.61.1.0031
- KAHN, J. S. (1989). Culture: Demise or resurrection? *Critique of Anthropology*, 9(2), 5–25. DOI: 10.1177/0308275X890900202
- KALENDER, B., TREHUB, S. E., & SCHELLENBERG, E. G. (2013). Cross-cultural differences in meter perception. *Psychological Research*, 77, 196–203. DOI: 10.1007/s00426-012-0427-y
- KENDAL, J. R. (2011). *Cultural niche construction and human learning environments: Investigating sociocultural perspectives*. *Biological Theory*, 6, 241–250. DOI: 10.1007/s13752-012-0038-2
- KROEBER, A. L., & KLUCKHOHN, C. (1952). *Culture: A critical review of concepts and definitions*. New York: Vintage Books.
- KUBIK, G. (2010). *Theory of African music* (Vol. 2). Chicago studies in ethnomusicology. Chicago, IL: The University of Chicago Press.
- KUHL, P. K. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics*, 50, 93–107. DOI: 10.3758/BF03212211
- KUHL, P. K., WILLIAMS, K. A., LACERDA, F., STEVENS, K. N., & LINDBLOM, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255, 606–608. DOI: 10.1126/science.1736364
- KUPER, A. (2003). *Culture: The anthropologists’ account* (5th printing). Cambridge, MA: Harvard University Press. (Original work published 1999)
- KVIFTE, T. (2007). Categories and timing: On the perception of meter. *Ethnomusicology*, 51, 64–84.
- LALAND, K. N., & O’BRIEN, M. J. (2011). Cultural niche construction: An introduction. *Biological Theory*, 6, 191–202. DOI: 10.1007/s13752-012-0026-6
- LARGE, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In S. Grondin (Ed.), *Psychology of time* (pp. 189–232). Bingley, UK: Emerald.
- LARGE, E. W., & JONES, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, 106, 119–159. DOI: 10.1037//0033-295X.106.1.119
- LARGE, E. W., & KOLEN, J. F. (1994). Resonance and the perception of musical meter. *Connection Science*, 6, 177–208. DOI: 10.1080/09540099408915723
- LARGE, E. W., & PALMER, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, 26, 1–37. DOI: 10.1207/s15516709cog2601\_1
- LARGE, E. W., & SNYDER, J. S. (2009). Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, 1169, 46–57. DOI: 10.1111/j.1749-6632.2009.04550.x
- LENTZ, C. (2016). Culture: The making, unmaking and remaking of an anthropological concept. *Working Papers of the Department of Anthropology and African Studies of the Johannes Gutenberg University Mainz*, 166. Retrieved from <http://www.ifeas.uni-mainz.de/Dateien/AP166.pdf>
- LERDAHL, F., & JACKENDOFF, R. (1983). *A generative theory of tonal music*. Cambridge, MA: MIT Press.
- LIVELY, S. E., LOGAN, J. S., & PISONI, D. B. (1993). Training Japanese listeners to recognize English /r/ and /l/: II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, 94, 1242–1255. DOI: 10.1121/1.408177
- LIVELY, S. E., & PISONI, D. B. (1997). On prototypes and phonetic categories: A critical assessment of the perceptual magnet effect in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 1665–1679.
- LIVELY, S. E., PISONI, D. B., YAMADA, R. A., TOHKURA, Y., & YAMADA, T. (1994). Training Japanese listeners to identify English /r/ and /l/: III: Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America*, 96, 2076–2087. DOI: 10.1121/1.410149
- LONDON, J. (2002). Cognitive constraints on metric systems: Some observations and hypotheses. *Music Perception*, 19, 529–550. DOI: 10.1525/mp.2002.19.4.529

- LONDON, J. (2012). *Hearing in time: Psychological aspects of musical meter* (2nd ed.). Oxford, UK: Oxford University Press.
- LONDON, J., HIMBERG, T., & CROSS, I. (2009). The effect of structural and performance factors in the perception of anacrusis. *Music Perception*, 27, 103–120. DOI: 10.1525/mp.2009.27.2.103
- LONGUET-HIGGINS, C., & LEE, C. S. (1982). The perception of musical rhythms. *Perception*, 11, 115–128. DOI: 10.1068/p110115
- MADISON, G., & MERKER, B. (2002). On the limits of anisochrony in pulse attribution. *Psychological Research*, 66, 201–207. DOI: 10.1007/s00426-001-0085-y
- MANNING, F. C., & SCHUTZ, M. (2016). Trained to keep a beat: Movement-related enhancements to timing perception in percussionists and non-percussionists. *Psychological Research*, 80, 532–542. DOI: 10.1007/s00426-015-0678-5
- MARCUS, S. L. (2007). *Music in Egypt: Experiencing music, expressing culture*. New York: Oxford University Press.
- MARTENS, P. A. (2011). The ambiguous tactus: Tempo, subdivision benefit, and three listener strategies. *Music Perception*, 28, 433–448. DOI: 10.1525/mp.2011.28.5.433
- MATTHEWS, T. E., THIBODEAU, J. N. L., GUNTHER, B. P., & PENHUNE, V. B. (2016). The impact of instrument-specific musical training on rhythm perception and production. *Frontiers in Psychology*, 7, 69. DOI: 10.3389/fpsyg.2016.00069
- MENARY, R. (2015). The aesthetic niche. *The British Journal of Aesthetics*, 54, 471–475. DOI: 10.1093/aesthj/ayu045
- MERKER, B., MADISON, G., & ECKERDAL, P. (2009). On the role and origin of isochrony in human rhythmic entrainment. *Cortex*, 45, 4–17. DOI: 10.1016/j.cortex.2008.06.011
- MOELANTS, D. (2006). Perception and performance of *aksak* metres. *Musicae Scientiae*, 10, 147–172. DOI: 10.1177/102986490601000201
- NAVEDA, L., GOUYON, F., GUEDES, C., & LEMAN, M. (2011). Microtiming patterns and interactions with musical properties in samba music. *Journal of New Music Research*, 40, 225–238. DOI: 10.1080/09298215.2011.603833
- NEUHOF, H., POLAK, R., & FISCHINGER, T. (2017). Perception and evaluation of timing patterns in drum ensemble music from Mali. *Music Perception*, 34, 438–451. DOI: 10.1525/mp.2017.34.4.438
- NORENZAYAN, A., & HEINE, S. J. (2005). Psychological universals: What are they and how can we know? *Psychological Bulletin*, 131, 763–784. DOI: 10.1037/0033-2909.131.5.763
- PARNCUTT, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythm. *Music Perception*, 11, 409–464. DOI: 10.2307/40285633
- POLAK, R. (2010). Rhythmic feel as meter: Non-isochronous beat subdivision in jembe music from Mali. *Music Theory Online*, 16(4).
- POLAK, R. (2017). The lower limit for meter in dance drumming from West Africa. *Empirical Musicology Review*, 12, 205–226.
- POLAK, R., JACOBY, N., & LONDON, J. (2016). Both isochronous and non-isochronous metrical subdivision afford precise and stable ensemble entrainment: A corpus study of Malian jembe drumming. *Frontiers in Neuroscience*, 10, 285. DOI: 10.3389/fnins.2016.00285
- POLAK, R., & LONDON, J. (2014). Timing and meter in Mande drumming from Mali. *Music Theory Online*, 20(1).
- PORTERA, M. (2016). Why do human perceptions of beauty change? The construction of the aesthetic niche. *RCC Perspectives: Transformations in Environment and Society*, 5, 41–48. DOI: 10.5282/rcc/7728
- POVEL, D.-J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 3–18. DOI: 10.1037/0096-1523.7.1.3
- POVEL, D.-J. (1984). A theoretical framework for rhythm perception. *Psychological Research*, 45, 315–337. DOI: 10.1007/BF00309709
- RAVIGNANI, A., DELGADO, T., & KIRBY, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nature Human Behaviour*, 1(1), 7. DOI: 10.1038/s41562-016-0007
- RAVIGNANI, A., & MADISON, G. (2017). The paradox of isochrony in the evolution of human rhythm. *Frontiers in Psychology*, 8, 1820. DOI: 10.3389/fpsyg.2017.01820
- REPP, B. H. (1984). Categorical perception: Issues, methods, findings. *Speech and Language: Advances in Basic Research and Practice*, 10, 243–335. DOI: 10.1016/B978-0-12-608610-2.50012-1
- REPP, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *Journal of Motor Behavior*, 35, 355–370. DOI: 10.1080/00222890309603156
- REPP, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review*, 12, 969–992. DOI: 10.3758/Bf03206433
- REPP, B. H. (2008). Metrical subdivision results in subjective slowing of the beat. *Music Perception*, 26, 19–39. DOI: 10.1525/Mp.2008.26.1.19
- REPP, B. H. (2010). Sensorimotor synchronization and perception of timing: Effects of music training and task experience. *Human Movement Science*, 29, 200–213. DOI: 10.1016/j.humov.2009.08.002
- REPP, B. H., LONDON, J., & KELLER, P. E. (2005). Production and synchronization of uneven rhythms at fast tempi. *Music Perception*, 23, 61–78. DOI: 10.1525/mp.2005.23.1.61
- REPP, B. H., LONDON, J., & KELLER, P. E. (2011). Perception-production relationships and phase correction in synchronization with two-interval rhythms. *Psychological Research*, 75, 227–242. DOI: 10.1007/s00426-010-0301-8

- REPP, B. H., LONDON, J., & KELLER, P. E. (2012). Distortions in reproduction of two-interval rhythms: When the “attractor ratio” is not exactly 1:2. *Music Perception*, 30, 205–223. DOI: 10.1525/mp.2012.30.2.205
- RICHERSON, P. J., & BOYD, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago, IL: University of Chicago Press.
- RIEDE, F. (2011). Adaptation and niche construction in human prehistory: A case study from the southern Scandinavian Late Glacial. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366, 793–808. DOI: 10.1098/rstb.2010.0266
- ROSCH, E. (1975). Cognitive reference points. *Cognitive Psychology*, 7, 532–547. DOI: 10.1016/0010-0285(75)90021-3
- SADAKATA, M., DESAIN, P., & HONING, H. (2006). The Bayesian way to relate rhythm perception and production. *Music Perception*, 23, 269–288. DOI: 10.1525/Mp.2006.23.3.269
- SAMUEL, A. G. (1982). Phonetic prototypes. *Perception and Psychophysics*, 31, 307–314. DOI: 10.3758/BF03202653
- SAVAGE, P. E., BROWN, S., SAKAI, E., & CURRIE, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proceedings of the National Academy of Sciences*, 112, 8987–8992. DOI: 10.1073/pnas.1414495112
- SCHULTZ, E. A. (2014). New perspectives on organism–environment interactions in anthropology. In G. Barker, E. Desjardins, & T. Pearce (Eds.), *History, philosophy and theory of the life sciences* (Vol. 4, pp. 79–102). Dordrecht: Springer. DOI: 10.1007/978-94-007-7067-6\_5
- SCHULTZ, E. A. (2016). Niche construction and the study of culture change in anthropology: Challenges and prospects. *St. Cloud State University Anthropology Faculty Publications*, 3. Retrieved from [http://repository.stcloudstate.edu/anth\\_facpubs/3](http://repository.stcloudstate.edu/anth_facpubs/3)
- SCHULZE, H.-H. (1989). Categorical perception of rhythmic patterns. *Psychological Research*, 51, 10–15. DOI: 10.1007/Bf00309270
- SEMJEN, A., & IVRY, R. B. (2001). The coupled oscillator model of between-hand coordination in alternate-hand tapping: A reappraisal. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 251–265. DOI: 10.1037//0096-1523.27.2.251
- SHWEDER, R. A. (2010). Donald Campbell’s doubt: Cultural difference or failure of communication? *Behavioral and Brain Sciences*, 33, 109–110. DOI: 10.1017/S0140525X10000245
- SNYDER, J. S., HANNON, E. E., LARGE, E. W., & CHRISTIANSEN, M. H. (2006). Synchronization and continuation tapping to complex meters. *Music Perception*, 24, 135–146. DOI: 10.1525/mp.2006.24.2.135
- STEVENS, C. J. (2012). Music perception and cognition: A review of recent cross-cultural research. *Topics in Cognitive Science*, 4, 653–667. DOI: 10.1111/j.1756-8765.2012.01215.x
- STEVENS, C. J., & BYRON, T. (2016). Universals in music processing: Entrainment, acquiring expectations, and learning. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 19–31). Oxford, UK: Oxford University Press.
- STOBART, H., & CROSS, I. (2000). The Andean anacrusis? Rhythmic structure and perception in Easter songs of Northern Potosí, Bolivia. *British Journal of Ethnomusicology*, 9(2), 63–94. DOI: 10.1080/09681220008567301
- STOTZ, K. (2010). Human nature and cognitive–developmental niche construction. *Phenomenology and the Cognitive Sciences*, 9, 483–501. DOI: 10.1007/s11097-010-9178-7
- SUMMERS, J. J., BELL, R., & BURNS, B. D. (1989). Perceptual and motor factors in the imitation of simple temporal patterns. *Psychological Research*, 51, 23–27. DOI: 10.1007/BF00309272
- SUMMERS, J. J., HAWKINS, S. R., & MAYERS, H. (1986). Imitation and production of interval ratios. *Perception and Psychophysics*, 39, 437–444. DOI: 10.3758/BF03207072
- TAL, I., LARGE, E. W., RABINOVITCH, E., WEI, Y., SCHROEDER, C. E., POEPPPEL, D., & ZION GOLUMBIC, E. (2017). Neural Entrainment to the beat: The “missing-pulse” phenomenon. *The Journal of Neuroscience*, 37, 6331–6341. DOI: 10.1523/JNEUROSCI.2500-16.2017
- TENNIE, C., CALL, J., & TOMASELLO, M. (2009). Ratcheting up the ratchet: On the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364, 2405–2415. DOI: 10.1098/rstb.2009.0052
- TOIVIAINEN, P., & EEROLA, T. (2003). Where is the beat? Comparison of Finnish and South African listeners. In R. Kopiez (Ed.), *Proceedings of the 5th triennial conference of the European Society for the Cognitive Sciences of Music (ESCOM)* (pp. 501–504). Hanover, Germany: Institute for Research in Music Education.
- TOMASELLO, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- TOMASELLO, M., KRUGER, A. C., & RATNER, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, 16, 495–511. DOI: 10.1017/S0140525X0003123X
- TREHUB, S. E., BECKER, J., & MORLEY, I. (2015). Cross-cultural perspectives on music and musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370, 20140096. DOI: 10.1098/rstb.2014.0096
- TYLOR, E. B. (1920). *Primitive culture: Researches into the development of mythology, philosophy, religion, art, and custom* (6th ed.). London, UK: John Murray. (Original work published 1871)
- ULLAL-GUPTA, S., HANNON, E. E., & SNYDER, J. S. (2014). Tapping to a slow tempo in the presence of simple and complex meters reveals experience-specific biases for processing music. *PLoS ONE*, 9(7), e102962. DOI: 10.1371/journal.pone.0102962

WHITEHOUSE, H., & COHEN, E. (2012). Seeking a rapprochement between anthropology and the cognitive sciences: A problem-driven approach. *Topics in Cognitive Science*, 4, 404–412. DOI: 10.1111/j.1756-8765.2012.01203.x

WILL, U. (2011). Prospects for a reorientation in cognitive ethnomusicology. In W. Steinbeck & R. Schumacher (Eds.), *Kölner Beiträge zur Musikwissenschaft: Bd. 16. Selbstreflexion in der Musik/Wissenschaft. Referate des Kölner Symposions 2007: Im Gedenken an Rüdiger Schumacher* (pp. 193–211). Kassel, Germany: Gustav Bosse Verlag.

WILL, U. (2017). Cultural factors in responses to rhythmic stimuli. In J. R. Evans & R. P. Turner (Eds.), *Rhythmic stimulation procedures in neuromodulation* (pp. 279–306). London, UK: Academic Press. DOI: 10.1016/B978-0-12-803726-3.00009-2

YATES, C. M., JUSTUS, T., ATALAY, N. B., MERT, N., & TREHUB, S. E. (2017). Effects of musical training and culture on meter perception. *Psychology of Music*, 45, 231–245. DOI: 10.1177/0305735616657407

## Appendix: Nonmusicians

TABLE A1. *Nonmusician Groups Used for Comparison*

Country of Residence	Group	Number and gender	Age
Mali	University students with little experience in folk dance	$n = 11$ , 3 females	mean = 25, range = 19–35
Mali	Members of folk dance associations	$n = 12$ , 10 females	mean = 36, range = 13–52
Bulgaria	Members of folk dance associations	$n = 14$ , 11 females	mean = 34, range = 18–56
Germany	Music listeners with relatively little experience playing instruments	$n = 21$ , 15 females	mean = 33, range = 19–69
Total		$N = 58$ , 38 females	mean = 32, range = 13–69

### PARTICIPANTS

For control reasons, we included groups of “nonmusicians” whose tapping we tested only at the slow tempo (periodicity = 1000 ms). The nonmusicians were dancers from Mali and Bulgaria familiar with the styles of dance associated with music played by our Malian drummer and Bulgarian folk musician groups, respectively, as well as university students (non-music majors and young graduates) and general population from Mali and Germany. Members of these groups have relatively little experience with playing instruments and public performances and would not usually identify as musicians in their societal contexts.

### RESULTS

Differences between musician and nonmusician groups emerged during efforts to pilot the experiment in Germany and Mali, which preceded the main experiments reported in this paper. Whereas both groups were able to perform the various tapping tasks at the slower tempo (periodicity of 1000 ms), many nonmusicians had difficulty performing the tapping task at the faster tempo (periodicity of 500 ms). For this reason, we did not try to collect data from nonmusicians at the faster tempo.

The cluster of nonmusician groups was less consistent than the cluster of musician groups in their isochronous tapping at the slow tempo (IOI = 500 ms), as indicated by the larger standard deviations of their

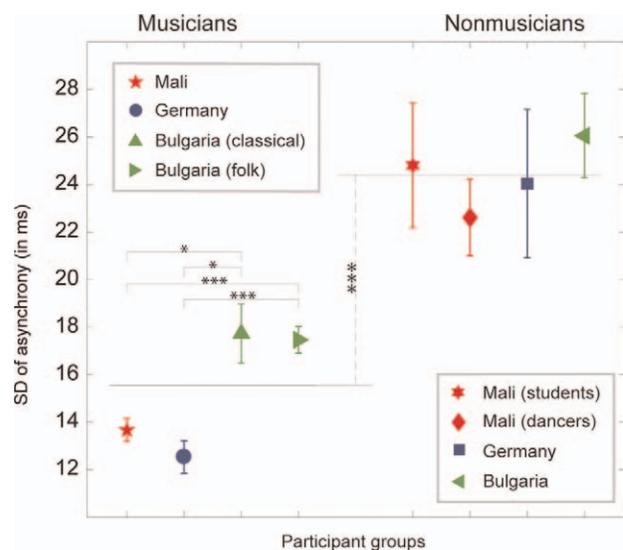


FIGURE A1. Consistency of synchronization (standard deviation of asynchronies) between stimulus and response onsets with an isochronous (1:1) target rhythm at the slow tempo, separated by “musician” versus “nonmusician” clusters of participant groups. The tested groups are plotted to the x-axis; the y-axis shows the mean values (symbol) and standard errors (error bars) of the standard deviation in milliseconds. Grey horizontal lines illustrate the grand average values for the musician and nonmusician clusters of groups, respectively. \*\*\* =  $p < .001$ ; \* =  $p < .05$ ; all  $p$  values were corrected for multiple comparison with Bonferroni correction.

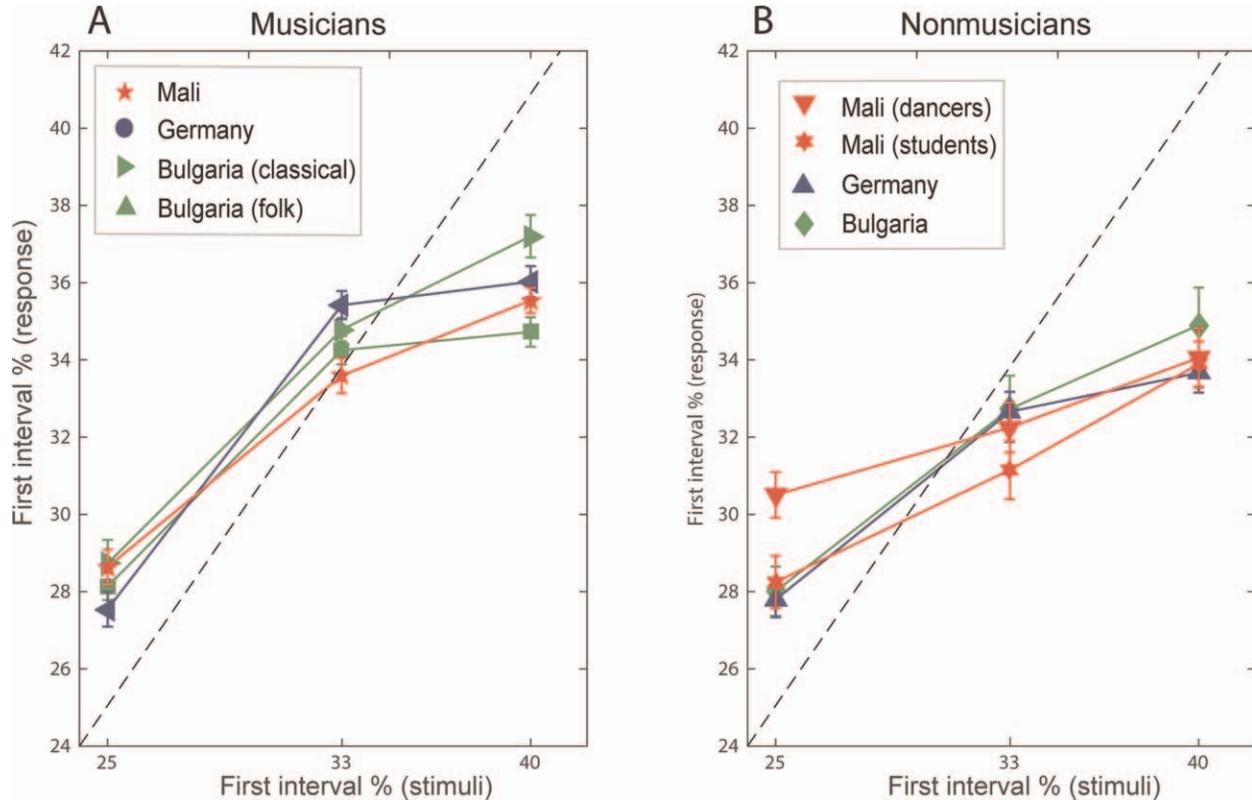


FIGURE A2. Responses to the 3:1, 2:1, and 3:2 ratios at the slow tempo, separated by expertise clusters (musician vs. nonmusician groups).

asynchronies from the target ratios; see Figure A1;  $t(126) = -6.77, p < .001$ . Remember that our periodicity, tempo, and event-rate calculations all refer to two-element rhythms. Thus, the pattern periodicity of 1000 ms in the slow tempo gives a rate of 500 ms when the ratio of the two events is 1:1. Note that there are some small but statistically significant differences among the musician groups, which are marked in Figure A1. The two groups of Bulgarian musicians showed a somewhat lesser degree of consistency than both Malian and German musicians,  $t(68) = 6.26, p < .001$ . The absolute value of this difference is quite small, however

(approximately 4 ms). By contrast, there is a considerably larger difference (approximately 9 ms) between musician and nonmusicians clusters.

We also found consistent differences between musician and nonmusician group clusters in the ratios of the reproduction of the 2:1 target at the slow tempo. While the musician groups slightly yet significantly softened the ratio,  $t(69) = -5.02, p < .001$ , toward approximately 1.9:1, the nonmusician groups instead sharpened the 2:1 ratio to approximately 2.1:1 (see Figure A2). This difference between musician and nonmusician groups was significant,  $t(126) = -5.41, p < .001$ .