



ARTICLE

<https://doi.org/10.1057/s41599-019-0363-1>

OPEN

What the music said: narrative listening across cultures

Elizabeth Hellmuth Margulis^{1*}, Patrick C.M. Wong², Rhimmon Simchy-Gross³ & J. Devin McAuley^{4*}

ABSTRACT Instrumental music can seem to tell engrossing stories without the use of words, but it is unclear what leads to this narrativization. Although past work has investigated narrative responses to abstract moving shapes, very little work has studied the emergence of narrative perceptions in response to nonlinguistic sound. We measured narrative responses to wordless Western and Chinese music in participants in the US and in a cluster of villages in a rural part of China using a Narrative Engagement (NE) scale developed specifically for this project. Despite profound differences in media exposure, musical habits, and narrative traditions, narrative listening was employed by many participants and associated with enjoyment in both groups; however, the excerpts that unleashed this response were culture-specific. We show that wordless sound is capable of triggering perceived narratives in two groups of listeners with highly distinct patterns of cultural exposure, reinforcing the notion that narrativization itself is a readily available mode of experiencing music. The particular sounds that trigger narrativization, however, rely on enculturation processes, as demonstrated by the within-culture consistency, but between-culture divergence in the specific excerpts that led to narrative engagement. Narratives can emerge in multiple modalities, including wordless sound, but association patterns specific to individual cultures critically shape how apparently abstract sound patterns come to acquire deep meaning and significance to people.

¹Department of Music, Princeton University, Woolworth Center of Musical Studies, Princeton, NJ 08544-1007, USA. ²Brain and Mind Institute, Department of Linguistics and Modern Languages, The Chinese University of Hong Kong, G17, Leung Kau Kui Building, Hong Kong SAR, China. ³Department of Psychology, Missouri Western State University, Murphy Hall 217, Saint Joseph, MO 64507, USA. ⁴Department of Psychology, Michigan State University, 316 Physics Rd., Room 262, East Lansing, MI 48824, USA. *email: margulis@princeton.edu; dmcauley@msu.edu

Introduction

Humans possess a robust tendency to narrativize abstract events. In the 1940s, Heider and Simmel (1944) conducted a simple experiment in which they showed people a film clip of shapes moving around a screen. When participants were asked to describe what they had seen, they tended to view the shapes as interacting animate beings with agency, independent of instructions. To make sense of the apparently abstract geometric shapes and their movements, people inferred a narrative. According to its sparest definition, narrative entails the recounting of an event or a series of events (Abbott, 2008). The impulse to narrativize has been characterized as distinctively human, shaping the way people make sense of their lives (McAdams, 2006). In the domain of music, people's tendency to narrativize, even for music without lyrics, is well-illustrated by the literary scene from *Howards End*, where Helen hears "heroes and shipwrecks" listening to Beethoven's Fifth Symphony (Forster, 1910, p. 30). This literary example concisely illustrates the idea that people experience music not just as abstract sequences of sound, but—often—as episodes in an unfolding narrative. What makes people hear music this way? How are their narratives shaped by patterns in the musical piece and by the culture in which they live?

Scientific work on music and narrative processing is minimal. In fact, a foundational axiom of much work in the psychology of music has been that music is a kind of abstract stimulus, not mired in referential content like language. Besson and Friederici (1998, p. 5) articulate this assumption: "One of the clearest differences between language and music is that whereas language is understood by reference to an extralinguistic designated space, music is self-referential. Thus, whereas words have meaning by convention, notes or chords have meaning by reference to previous notes." Yet a recent study tracked narrative responses to music for a small number of excerpts by participants with broadly shared cultural backgrounds and found not only that people frequently reported hearing wordless excerpts in terms of a perceived story, but also that participants generated remarkably consistent stories for each excerpt (Margulis, 2017)—one seemed to the majority of listeners like a cat and mouse chase, for example, and another like a soiree at a fancy ballroom. Enculturation—an individual's gradual adaptation to the characteristics of a culture through short-term or long-term exposure—likely influences whether and the extent to which listeners structure what they hear as a narrative. Repeatedly encountering particular sound patterns within specific contexts might foster concrete associations (e.g. a national anthem calling to mind an image of a country's flag, or the screeching, high-pitch sound of a violin evoking the shower scene from *Psycho*) that influence narrative engagement (NE) for enculturated listeners. Indeed, Tagg and Clarida (2003) and Huovinen and Kaila (2015) played people excerpts of film, television, and production music and found they were able to reliably predict what scenes might accompany it.

Since notions of story structure vary interculturally (Herman et al., 2012), as do musical systems (Gourlay, 1984), it is uncertain how the phenomenon of narrative listening might extend across cultures. If—following Heider and Simmel's findings—humans possess a fundamental tendency to understand abstract stimuli in terms of narrative, narrative listening to music might arise to the same degree and in similar ways across cultures. Otherwise, narrative listening might differ in extent and kind when people with different backgrounds and musical exposure patterns are examined. A number of previous studies have revealed the impact of culture and experience on aspects of music perception (Hannon and Trehub, 2005; Wong et al., 2011), constituting diverse examples of the broader phenomenon of perceptual learning. This article questions culture's impact on whether or not people

listen to music narratively in the first place, targeting not just the sensory details of perception, but the actual existence (or not) of an actual *mode* of perception—namely, narrative listening.

To address the phenomenon of narrative listening and the role culture plays in its deployment, narrative responses to music were compared for two culturally distinct participant groups—one comprised of students in the midsouth of the United States (the Arkansas participant group) and the other one comprised of residents of a remote cluster of Dong villages in Guizhou, China (the Dimen participant group). People in Dimen speak a variety of Dong, a tone language that belongs to the Tai family (Ramsey, 1989, p. 230), which is independent of languages within the Sino-Tibetan family—including Mandarin—spoken by the majority of the population in China. Dong possesses no widely used written form, and most participants in the study were not literate. Although people in Dimen have exposure to popular Chinese media, they have very little exposure to Western media and their proficiency in (spoken) Mandarin and written Chinese is limited or absent. Thus, participants in the Dimen group possess very different framing experiences with narrative. Whereas participants in the Arkansas group grew up with access to stories that had been fixed into written form, participants in the Dimen group grew up experiencing stories within an exclusively oral tradition. Thus, their basic conceptualizations of narrativity likely differed substantially, rendering any similarities in their deployment of the notion during musical listening more striking.

A third participant group in the US Midwest (the Michigan participant group) was included as part of the study to assess the within-culture consistency of responses from the Arkansas group. Importantly, none of these groups habitually listen to music without words. Participants in the Arkansas and Michigan groups reported listening most frequently to pop and rock songs and possessed little experience with the kind of Chinese-language media that can powerfully shape musical associations with Chinese music. Participants in the Dimen group reported listening most frequently to Big Song, a polyphonic singing tradition unique to their minority group—one of 55 officially recognized minority groups in China and conversely possessed little experience with the kind of English-language media that can shape musical associations with Western music.

Narrative Engagement (NE) with music was assessed for Arkansas, Dimen, and Michigan participants using both a direct question about whether or not participants imagined a story while listening (the Story Response Question or SRQ), and a specially devised NE scale, detailed in the "Materials and methods" section. Stimuli were 128 60-s excerpts of wordless music, half Chinese and half Western. In both cases, they were drawn from styles of music with which participants did not habitually engage—specifically, Western classical music and Chinese art music. This ensured that participants' responses did not reflect explicit instruction in appropriate ways to hear these styles. Instead, previous experiences with the musical styles used in the experiment transpired in more ambient, passive contexts, such as media consumption. Although there were significant cultural differences between participants in the Arkansas and Michigan groups and the Dimen group, a parallel design was used.

Materials and methods

Materials

Stimuli. Stimuli were 60 s excerpts ($n = 128$) drawn from commercial recordings of instrumental music with no lyrics or vocal part. Half of the excerpts ($n = 64$) were drawn from recordings of Western art music and half ($n = 64$) from recordings of Chinese art music. Preliminary pilot work determined that although

participants in the Dimen group were broadly familiar with the style of Chinese music presented in the experiment, and participants in the Arkansas and Michigan groups were broadly familiar with the style of Western music presented in the experiment, these styles of music were not the ones to which participants tended to listen most frequently, and these specific excerpts were unlikely to be ones that participants had heard prior to the experimental session (i.e., they were relatively novel).

NE scale. Authors JDM and EHM and a group of eight subject matter experts (comprised of members of the Digital Humanities and Literary Cognition lab at MSU who were advanced undergraduates or graduate students in English) generated a set of potential survey items. Candidate items were either generated through an initial literature review or were brainstormed based on subject-matter expertise in NE.

Next, candidate items that were determined to be awkwardly or ambiguously phrased by the subject matter experts were eliminated, resulting in a list of 38 items that targeted NE. An initial sample of 202 participants ($n = 155$, female), aged 18–33 years ($M = 20.3$, $SD = 1.6$) at Michigan State University listened to a 5-min instrumental musical excerpt that was from a Western art music tradition, and then completed the initial 38-item survey. In addition to completing the survey, participants were asked how familiar and enjoyable they found the excerpt and if they imagined a story or elements of a story when listening to the excerpt. If they did imagine a story or elements of a story, they then provided a short description of the story they imagined. An initial exploratory factor analysis on the responses to the 38-item survey using PCA led to the identification of a single NE factor that included 12 items with a loading of >0.70 .

To examine whether the survey results from the Western musical excerpt generalized to a Chinese musical excerpt, a second sample of 126 participants ($n = 82$ female) at Michigan State University, aged 18–52 years ($M = 20.2$, $SD = 3.6$), listened to a 5-min instrumental musical excerpt drawn from a Chinese art music tradition and completed the same task. A second exploratory factor analysis similarly identified a single NE factor; 8 of the 12 items with factor loadings that had been >0.70 for the initial sample (drawn from Western art music) again demonstrated factor loadings >0.70 for the second sample (drawn from Chinese art music). From these eight items, we selected four items for an abbreviated version of the NE scale; the wording of one of the items was slightly revised to improve clarity. The four NE items were ‘It was easy to imagine a story when listening to the music’, ‘I imagined a vivid story’, ‘I imagined a story with a clear setting, characters, and events’, and ‘I imagined a story while the music was playing, not afterwards’. Cronbach alpha for the final four-item narrative (NE) scale indicated good internal consistency for both the Western excerpt sample ($\alpha = 0.86$) and the Chinese excerpt sample ($\alpha = 0.89$).

Criterion validity was assessed by examining the relation between NE scores and characteristics of the story descriptions that participants wrote in response to the Western excerpt (sample 1) and the Chinese excerpt (sample 2). For both the Western and Chinese excerpts, higher NE scores were strongly associated with the number of events in the stories that participants generated (Western excerpt, $r = 0.47$, $p < 0.001$; Chinese excerpt, $r = 0.54$, $p < 0.001$), the number of characters in the story (Western excerpt, $r = 0.51$, $p < 0.001$; Chinese excerpt, $r = 0.47$, $p < 0.001$), the number of words in the story (Western excerpt, $r = 0.46$, $p < 0.001$; Chinese excerpt, $r = 0.51$, $p < 0.001$), and an independent (blind) raters’ assessment of story richness (Western excerpt, $r = 0.55$, $p < 0.001$; Chinese excerpt, $r = 0.58$, $p < 0.001$).

Experimental design. The general design was a 2 (Group: Arkansas, Dimen) \times 2 (culture-of-excerpt: Western, Chinese) mixed factorial. Each participant listened to a set of eight excerpts (four Western and four Chinese) from the pool of 128. For each excerpt, participants completed the NE scale, responded to items assessing enjoyment and familiarity, and provided either a description of the story they heard for each excerpt or a description of why they thought that they did not hear a story (these free-response descriptions are reported in a separate manuscript). Different participants heard different sets of eight excerpts, so that narrative responses were obtained for all 128 excerpts across all participants.

Participants. The Arkansas group consisted of 321 participants ($n = 197$ female, 121 male, 3 no response; ages 18–41 years, $M = 19.4$, $SD = 2.3$); participants took part in the study at the Music Cognition Lab at the University of Arkansas in Fayetteville, Arkansas. The Dimen group consisted of 141 participants (123 female, 18 male, 0 no response; ages 20–81 years, $M = 48.55$, $SD = 13.1$) who took part in the study at the Dimen Dong Community Cultural Research Center in Dimen, Guizhou Province, China. Regardless of their proficiency level in Chinese, all participants were consented with the presence of a Dong translator. The person obtaining consent spoke Mandarin to the participants and the translator, and explained the entire experimental process as well as the risks of the experiments to them. The participants who did not have sufficient literacy to write their name were instructed to draw a symbol on the consent form. 65 participants had sufficient written Chinese proficiency to sign their name. Even participants who could sign their name did not possess more than a few years of formal literacy instruction. A third smaller sample of 39 listeners ($n = 27$, 19–68 years, $M = 31.8$, $SD = 12.2$) from the East Lansing, MI, community (Michigan group) participated in the study to assess the within-culture consistency of the responses from the Arkansas group. Each Michigan participant heard a larger number of excerpts so that the number of responses per excerpt was approximately the same for the Michigan and Arkansas groups. Similar to the Arkansas group, participants in the Michigan group varied in their formal music training (0–25 years, $M = 3.0$, $SD = 4.9$).

Procedure. Once seated for the experiment, participants were instructed “You’ll be asked to report aspects of your experience listening to musical excerpts, including whether or not you imagined a story while listening. Please do NOT specifically ATTEMPT to imagine a story. Simply listen to the music as you ordinarily would. If you imagine a story, that’s fine, and if you don’t imagine a story, that’s fine too.” Following these instructions, each participant heard a subset of eight musical excerpts from the full set of 128—four Western and four Chinese. Each excerpt was followed by a series of questions, and the excerpt order was randomized for each participant. Prior to each excerpt, participants were told they should try to listen attentively as if they were intending to enjoy the piece. After listening to each excerpt, participants indicated whether they imagined a story or elements of a story while listening to the music (yes/no): the SRQ.

Next, they completed a four-item version of the NE scale. For the NE scale items, participants were asked to rate the extent to which they endorsed a statement about the music ranging from 1—strongly disagree to 6—strongly agree. Finally, they indicated whether they had heard the specific excerpt before, whether it sounded familiar to them, and whether they enjoyed listening to the excerpt using the same 6-point rating scale. After completing these items, participants responded to one of two free-response

questions, depending on their response to the SRQ. If they answered yes to the SRQ, they described the story they imagined in as much detail as they were able. If they answered no to the SRQ, they were asked to speculate about why they did not imagine a story. Requesting free responses in both cases ensured that participants were not incentivized to select one option or the other simply on the basis of extent of subsequent task demands.

Participants in the study listened to the excerpts over high-quality headphones. Participants at the Arkansas and Michigan research sites completed the task in English, entering responses directly into the computer. Researchers administering the study at the Dimen site were aided by an experienced translator who was fluent in Dong and Mandarin. Participants at the Dimen research site completed the task in their own language, with research assistants recording their spoken responses, which were then transcribed. At all three performance sites, participants were encouraged to ask questions if any instructions were unclear and data collection did not start until the researchers, including the translator in the case of the Dimen participants, were confident that participants understood the task. The entire experiment took ~50 min.

Statistical analysis. Analyses focused on the quantitative data, with qualitative free-response data reported in a separate paper. Ratings on a 6-point ratings scale (1—strongly disagree, 6—strongly agree) for the four items from the NE scale were averaged to create a composite NE score. Quantitative data were separately (1) combined across excerpt within each excerpt type category (Western or Chinese) and (2) combined across subjects in each group for each excerpt in order to permit both subject and item analyses of the data. Primary statistical comparisons were between the Arkansas and Dimen groups, with data from the Michigan group used to assess the consistency of responses by participants in the Arkansas group. Age, enjoyment, and familiarity ratings for participants in Arkansas and Dimen were included as covariates in all group comparisons using $\alpha = 0.05$ for statistical significance. Pearson correlation coefficients were calculated to examine relations between variables and within-culture and between-culture consistency of by-excerpt responses across groups. Correlational analyses used a Bonferroni correction to minimize family-wise error rate for each set of analyses using $\alpha_{FW} = 0.05$ for statistical significance.

Results

Figure 1 shows that narrative listening is a mode of musical response that is not restricted to Western, Industrialized, Rich and Democratic (WEIRD) societies (Henrich et al., 2010). Via answers to the SRQ, participants in Fayetteville, Arkansas indicated that they imagined a story ~66% of the time in response to any given excerpt, while participants in Dimen, China reported imagining a story ~50% of the time in response to any given excerpt. A two-way mixed-measures ANOVA on SRQ scores with participant age, and ratings of excerpt enjoyability and familiarity included as covariates, revealed that SRQ scores were greater for the Arkansas participants than for the Dimen participants, $F(1457) = 29.68$, $p < 0.001$, but there was also a reliable interaction between group and culture of excerpt, $F(1457) = 9.74$, $p = 0.02$ (see Fig. 1a).¹

Although participants narrativized in response to music from both cultures, they were significantly more likely to listen narratively to music with which they had more cultural exposure. Arkansas listeners were more likely to imagine a story for Western musical excerpts than Chinese musical excerpts (Western, $M = 71\%$, $SD = 24\%$; Chinese, $M = 62\%$, $SD = 27\%$), $t(320) = 4.78$, $p < 0.001$. Conversely, Dimen listeners were more

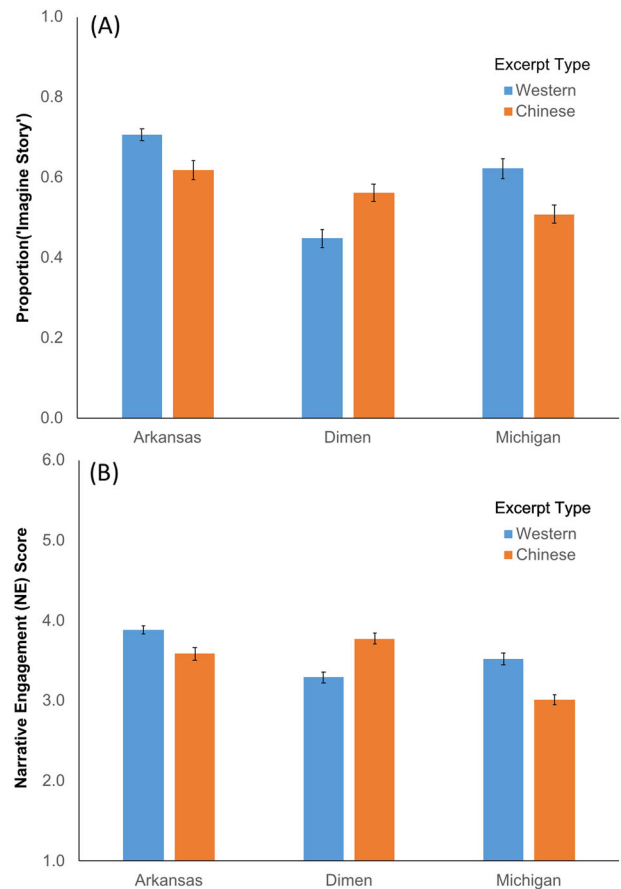


Fig. 1 Proportion 'yes' responses for the story response question (SRQ) and narrative engagement (NE) scores separated by excerpt type (Western vs. Chinese) for the three groups of listeners (Arkansas, Dimen, Michigan) are shown in panels **a**, **b**, respectively. Listeners from Arkansas and Michigan were more likely to imagine a story for Western excerpts than for Chinese excerpts, while listeners from Dimen were more likely to imagine a story for Chinese excerpts than for Western excerpts. NE scores followed a similar pattern for all three groups. Participants in the Arkansas and Michigan groups had higher NE scores for Western excerpts than for Chinese excerpts, while the reverse was true for the Dimen group

likely to imagine a story for Chinese musical excerpts than Western musical excerpts (Western, $M = 45\%$, $SD = 34\%$; Chinese, $M = 56\%$, $SD = 34\%$; $t(140) = -4.43$, $p < 0.001$). There was no overall main effect of the culture of the excerpt, $F(1457) = 0.15$, $p = 0.70$, indicating that neither type of excerpt was intrinsically more narrative, but that differences were driven by enculturation.

Analysis of the composite NE score that included four-items ('It was easy to imagine a story when listening to the music', 'I imagined a vivid story', 'I imagined a story with a clear setting, characters, and events', and 'I imagined a story while the music was playing, not afterwards') revealed a similar pattern (see Fig. 1b). A two-way mixed-measures ANOVA on NE scores with age, enjoyability, and familiarity as covariates revealed that NE was greater for Arkansas participants than for Dimen participants, $F(1457) = 8.34$, $p = 0.02$, but there was also a significant interaction between group and culture-of-excerpt, $F(1457) = 16.09$, $p < 0.001$. NE was higher for Arkansas participants for Western excerpts than for Chinese excerpts (Western, $M = 3.89$, $SD = 0.85$; Chinese, $M = 3.59$, $SD = 0.95$; $t(320) = 5.41$, $p < 0.001$), while the reverse was true for Dimen participants (Western, $M = 3.29$, $SD = 1.20$; Chinese, $M = 3.77$, $SD = 1.23$; $t(140) = 6.77$, $p < 0.001$).

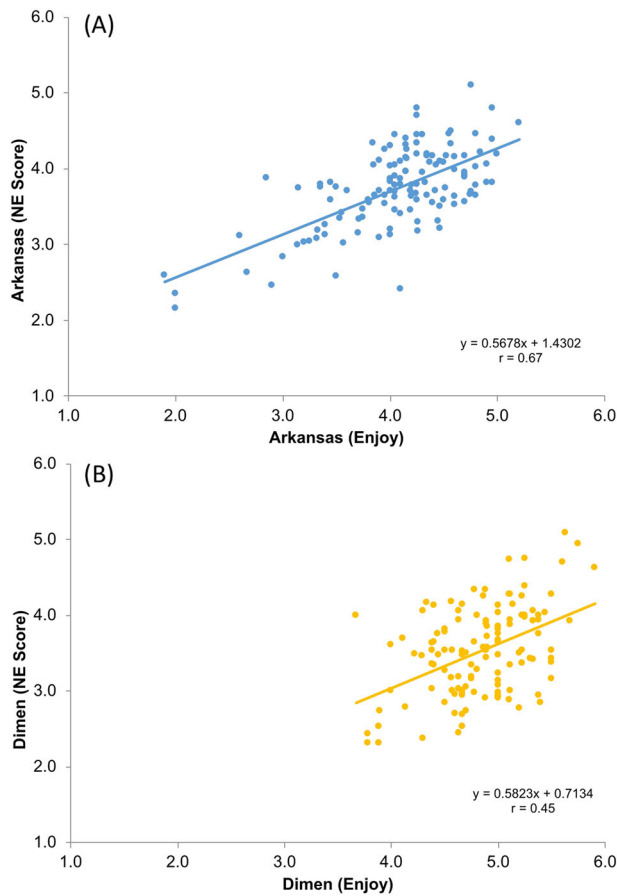


Fig. 2 Relation between enjoyment ratings and narrative engagement (NE) scores. Enjoyment ratings and NE scores were positively correlated for Arkansas participants **a** and Dimen participants **b**. The relation remained reliable after controlling for familiarity (Arkansas, $r = 0.45$, $p < 0.001$; Dimen, $r = 0.24$, $p < 0.01$)

Next, to examine the relation between NE and musical enjoyment, and to consider whether this association was culturally dependent, Pearson correlation coefficients were calculated between by-excerpt NE and enjoyment. Figure 2 shows that enjoyment was positively correlated with NE both for participants in Arkansas and participants in Dimen (Arkansas, $r = 0.66$, $p < 0.001$; Dimen, $r = 0.45$, $p < 0.01$). This relation tended to hold when data were separated by the excerpts' culture of origin. For the Arkansas group, correlations between enjoyment and NE for the Western and Chinese excerpts were $r = 0.32$, $p = 0.01$, and $r = 0.77$, $p < 0.001$, respectively. Similarly, for the Dimen group, correlations between enjoyment and narrative listening for the Western and Chinese excerpts were $r = 0.31$, $p = 0.01$, and $r = 0.45$, $p < 0.001$, respectively. The relation between enjoyment and NE further held for both groups when controlling for familiarity (Arkansas, $r = 0.45$, $p < 0.001$; Dimen, $r = 0.24$, $p < 0.01$).

The results for the SRQ showed the same pattern as the NE results. For the Arkansas group, the correlation between enjoyment and SRQ score was $r = 0.63$, $p < 0.001$. When the excerpts were separated by culture of origin, the positive correlation between enjoyment and proportion of 'yes' story responses was still reliable (Western: SRQ, $r = 0.26$, $p < 0.05$; Chinese: SRQ $r = 0.74$, $p < 0.001$). For the Dimen group, the correlation between enjoyment and proportion of 'yes' story responses was also significant, $r = 0.31$, $p < 0.01$. Similar to the Arkansas group, when the correlations by the excerpts were separated by culture of origin, the correlation between enjoyment and proportion of 'yes'

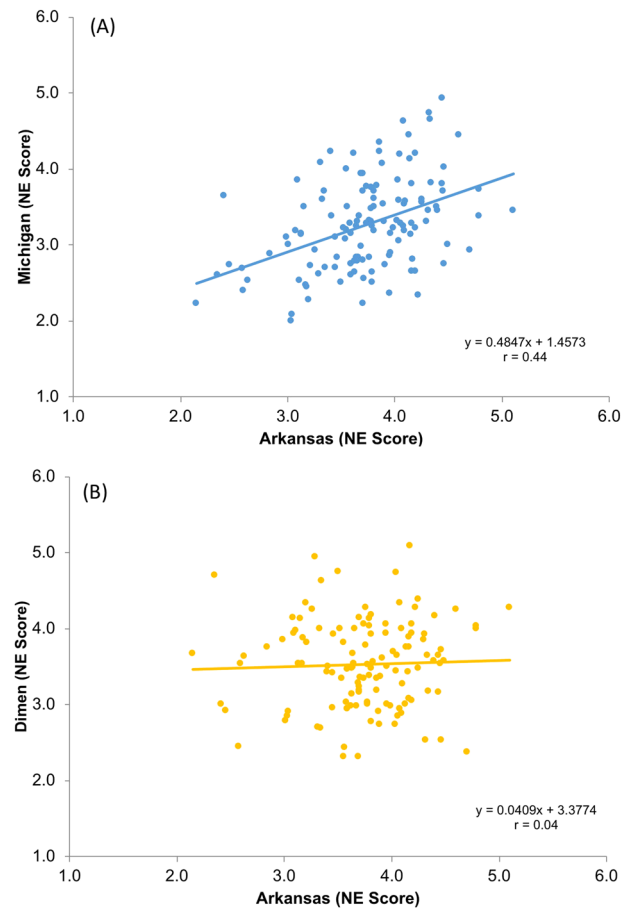


Fig. 3 Within-culture and between-culture comparison of narrative engagement (NE) by excerpt. NE scores were positively correlated for the within-culture comparison of the Arkansas and Michigan groups **a**, while there was no relation between NE scores for the between-culture comparison of the Arkansas and Dimen groups **b**

story responses was significant for the Chinese excerpts, but not for the Western excerpts (Western: SRQ, $r = 0.16$, $p = 0.22$; Chinese: SRQ = 0.33, $p < 0.01$).

Next, to assess whether NE depends exclusively on enculturated associative patterns, or whether it arises more spontaneously from abstract characteristics of the sound structure, a series of within-culture and between-culture by-excerpt correlations examined the consistency of listeners' responses both within and across cultures. Here, the data from the Michigan group was used to evaluate the within-culture consistency of the responses from the Arkansas group. Correlations by excerpt between the Arkansas and Michigan groups revealed a high degree of consistency in NE scores and enjoyment across excerpts (see Figs 3a and 4b). Correlations between the Arkansas and Michigan groups were positive both for NE scores ($r = 0.43$, $p < 0.001$) and for enjoyment ($r = 0.64$, $p < 0.001$).

A different picture emerged, however, for the between-culture comparisons between the Arkansas and Dimen groups (see Figs 3b and 4b). Although both the Arkansas and Dimen groups showed a high level of NE, the specific excerpts that led to a tendency to imagine a story, narratively engage with it, or enjoy it were unrelated. NE scores for individual excerpts were not correlated between the Arkansas and Dimen groups ($r = 0.04$, $p = 0.67$), or for the excerpts separated by culture of origin (Western: $r = 0.14$, $p = 0.27$; Chinese, $r = 0.21$, $p = 0.1$). The same picture emerged for enjoyment. Enjoyment ratings for individual excerpts were not

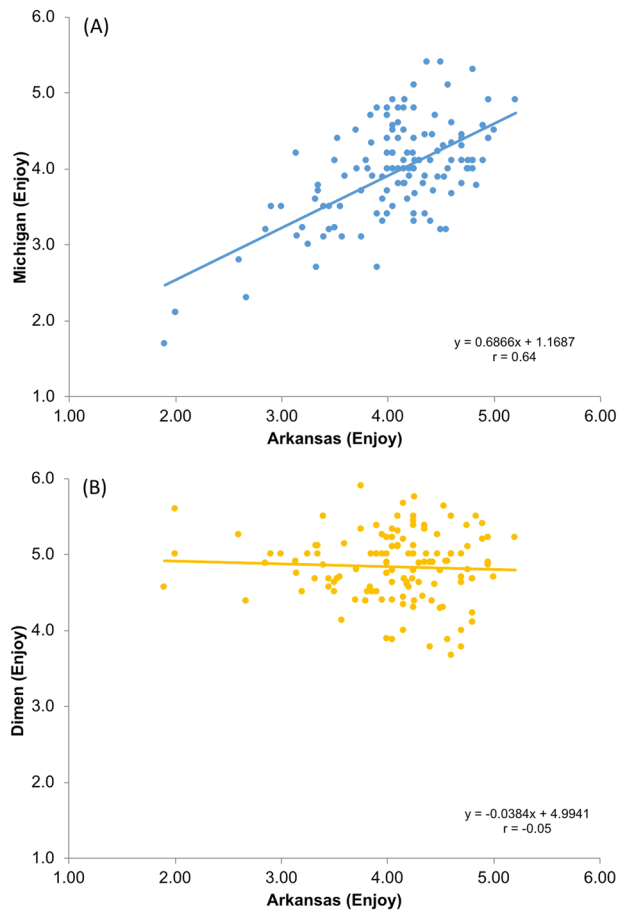


Fig. 4 Within-culture and between-culture comparison of enjoyment by excerpt. Enjoyment ratings were positively correlated for the within-culture comparison of the Arkansas and Michigan groups **a**, while there was no relation between the enjoyment ratings for the between-culture comparison of the Arkansas and Dimen groups **b**

correlated between the Arkansas and Dimen groups ($r = -0.05$, $p = 0.56$), or for the excerpts separated by culture of origin (Western: $r = 0.03$, $p = 0.81$; Chinese: $r = 0.025$, $p = 0.84$).

Similar to NE scores, there was a high-degree of within-culture consistency of the by-excerpt SRQ scores between the Arkansas and Michigan groups, $r = 0.30$, $p = 0.001$. In contrast, there was little between-culture consistency of the by-excerpt SRQ scores for the Arkansas and Dimen groups ($r = 0.12$, $p = 0.2$).

Finally, participants provided a free response description of all imagined stories. These descriptions revealed coherent stories featuring within-culture consensus akin to the rates in Margulis (2017); e.g. one excerpt was described by over 70% of respondents as expressing nature coming to life at sunrise, but another was consistently heard to imply a parade full of pomp. However, the content of the stories people generated differed, in some cases, dramatically, across cultures. For example, consider an atonal excerpt by composer Anton Webern. This excerpt thwarted the principles of Western tonality such that it was challenging to hear any one pitch as the central or referent note. In response, listeners in Arkansas reported stories featuring horror, murder, and paranoia. Listeners in Dimen, however, reported stories featuring playful and happy times with friends. Presumably, the lack of tonality had been the dominant musical attribute for listeners in Arkansas who were accustomed to hearing music with a Western tonal framework, but listeners in Dimen, not attempting to impose a tonal frame, were able to tune into the staccato

articulations, which could be read as playful. The degree to which descriptions cohere within cultures for individual excerpts but differ across excerpts demonstrates that the perceived stories are driven by the music, rather than arbitrarily paired with it. Moreover, scores on the NE scale correlate highly with the richness, structure, and detail of the free response stories, as described in the “Materials and methods” section.

Discussion

Narrative listening to wordless instrumental music was shown to be a readily available mode of response for many people in three samples of listeners, two from the Midwest and Midsouth of the United States, and one from a Dong community in a rural, mountainous part of Guizhou, China. On the one hand, this suggests that the capacity to experience music narratively is widespread, and not dependent on exposure to mass media. On the other hand, the specific excerpts that triggered this narrativization, while broadly the same for two sites in the central US, were completely unrelated between US sites and the location in rural China. This divergence suggests that particular mappings between sound pattern and story are highly dependent on enculturation. Together, these findings paint the picture that narrative listening is a fundamental mode of engaging with music, with enculturation determining which specific musical patterns seem to tell stories. The idea that narrative listening is one of a handful of basic strategies for making sense of music, but that cultural exposure shapes which stories develop in response to which music, helps illuminate why so many people seem tempted by the notion that music is a universal language. It is relatively easy to hear expressive and narrative characteristics in musical passages, making it seem as though these associations were universally present. But in reality, the specific implementation of these narrative associations depends critically on enculturation, ensuring that what sounds like a story about horror or murder to people in one place can sound like a happy story about friends to people in another, and more broadly that an excerpt that seems overwhelmingly narrative in character to one group of people can seem utterly free of narrative associations to people in another place.

We will first consider the finding that narrative listening seems to be a readily available mode of apprehending music. Across two cultures with very different musical traditions and narrative practice, when prompted about the possibility, many participants reported listening to wordless musical excerpts in terms of a story. Despite that participants in Dimen were steeped in a participatory singing tradition that differs substantially from the one to which the US participants had been exposed, both groups exhibited narrative listening as a mode of musical response, and were able to provide verbal descriptions of these stories with internal coherence. Although narrative listening was more common among US participants who have grown up in a culture saturated by the use of instrumental music in TV, film and video games, experiencing wordless musical sequences as the communication of a story extends beyond the specific case of participants who have grown up hearing instrumental music used this way. This discovery seems consistent with Heider and Simmel’s supposition that the tendency to narrativize abstract stimuli is broadly shared; however, the specific patterns of sound-narrative association vary with enculturation.

Overall, the results show that people’s capacity to narrativize in response to music does not depend on prior exposure to the specific sound sequences used, but rather on a presumed ability to abstract patterns and match them to the contexts in which similar patterns have appeared. Participants in both the US and China were able to listen narratively to excerpts they had never heard before the experimental session, and to listen narratively to excerpts in styles which they were less familiar. However, they were significantly more likely to narrativize in response to excerpts

in styles that were relatively more familiar (Chinese music for the Dimen listeners, Western music for the Arkansas and Michigan listeners). Consistent with this interpretation, the overall correlations between familiarity and NE were $r = 0.32$, $p < 0.001$, for the Arkansas group and $r = 0.79$, $p < 0.001$, for the Dimen group. Thus, the basic tendency to employ narrative listening is modulated by cultural exposure. Presumably, within-culture listeners possess more concrete mappings between sonic features and the contexts within which they are generally embedded.

The other major finding is that the individual excerpts that tended to elicit the narrative listening were different for participants in Dimen and at the US sites, suggesting that enculturated associative patterns drive the experience more than basic properties of the sound signal. Heider and Simmel noticed that when their film's shapes seemed to violate intuitive physics, by failing to conform to the movement patterns that would emerge if inertia or gravity were present, people were likelier to ascribe agency to the objects and understand them as interacting intentional agents. Similarly, music theorists have suggested that particular structural features trigger narrative listening, including melodic movements that violate intuitive physics or moments of noticeable contrast where the music's features change substantially (Almén, 2008). But if these or some other set of features inevitably encouraged narrativization, the same set of excerpts should elicit narrative listening at all sites. Because this was not the case, it seems unlikely that some simple, universal principle links acoustic features and perceived narrativization. Rather any structural affordances in the music were filtered through cultural associations to produce perceived stories.

Enjoyment and narrativization were correlated for participants at all sites—the excerpts that people narrativized to most strongly were also the ones they enjoyed the most. This was true despite the fact that the excerpts that were enjoyed and narrativized to at US sites were unrelated to the ones that were enjoyed and narrativized to at the site in China. This need not necessarily have been true; for example, narrativization could have been more effortful and unpleasant if narrativization were harder and less commonly employed within a particular cultural context. This robust correlation suggests that narrativization and enjoyment might be two frequent components of a broader engagement with music—a kind of attentive, interested, affectively absorbed mode of listening (Sloboda, 2011). When the music is engaging and enjoyable, it is easier to narrativize; similarly, when an excerpt can be heard narratively, it is engaging and enjoyable. The fact that these responses are intertwined supports the notion that narrativization is a fundamental form of positive engagement with music.

The ease with which people at all sites listened narratively suggests that existing work in music psychology on time estimation (Hui et al., 1997), emotional response (Juslin and Sloboda, 2011), preference (Rentfrow and Gosling, 2003), crossmodal associations (Vines et al., 2011), and beyond might be impacted by the mediating variable of narrativization, a response that heretofore has not been given significant attention—the target article of a major review of emotional responses to music (Juslin and Västfjäll, 2008), for example, contains no mention of it. These results also point to general strategies for the construction of meaning, relevance, and enjoyment in esthetic domains across cultures. Presented with largely unfamiliar music, participants at all sites used conceptual blending (Fauconnier and Turner, 2008) to understand it in terms of self-relevant stories. This speaks to the pervasiveness of storytelling as a mode for engaging with the world.

Scientists have long pursued potential overlaps between musical and linguistic processing. Processes involved in segmenting sound during music listening, for example, also appear to be involved in segmenting speech into meaningful units (Dilley and McAuley, 2008; Patel, 2008). Difficulties with rhythm processing, moreover, have been linked to a variety of developmental

language disorders, including dyslexia, stuttering, and specific-language impairment (Gordon et al., 2015; Wieland et al., 2015). Most directly relevant, Koelsch et al. (2004), in their EEG study of music, language, and meaning, found that tests of musical semantics via targeted priming elicited the same N400 component associated with linguistic processing, supporting the provocative notion that musical experiences may be mediated by verbal systems. The relationship this study documents between music and narrative processing strengthens the case that music perception draws on other cognitive resources, including linguistic ones, and not exclusively on specialized systems (Brown et al., 2006; Jackendoff and Lerdahl, 2006).

From a practical perspective, the correlation between narrativization and enjoyment suggests that outreach organizations and arts institutions aiming to generate increased engagement might focus on strategies that increase the likelihood and ease of narrative response. When relying on music to increase intercultural understanding, it is important to note that the excerpts that engender this response are not invariant across contexts. These are not magic, “universal language” excerpts that elicit narrative hearing in people regardless of culture. Instead, musical excerpts carry the residue and associations of the culture within which they are being heard, and activate the fundamental human capacity to narrativize differently depending on these background factors. Within a culture, however—even a broadly construed culture, such as “undergraduates at large public universities in the American midwest and midsouth”—people tend to narrativize to the same excerpts, exposing this tendency as an unconscious marker of group identity, similar to the dialect surveys that can identify a person's region of origin based on whether they say pop or soda for fizzy drinks and crawdad or crayfish for tiny crustaceans (Vaux and Golder, 2003). This notion, bolstered by the within-culture description consensus of the free response descriptions, warrants further investigation, especially as it intersects with music's critical role in identity formation (Garrido and Davidson, 2019).

Future studies should investigate narrativization to more excerpts in more cultures, including cultures that are defined in other ways than geography (Geertz, 1973). For example, the same paradigm could be employed with people who live in the same city but grew up in different decades or occupy different socioeconomic spheres. Subsequent work coming out of this project examines the commonalities and divergences within the narrative descriptions participants provide—a subject previous work has demonstrated can be characterized by a surprising amount of consensus on topic, structure, and even exact words used (Margulis, 2017). Other work pulls apart and recomposes individual excerpts in an attempt to pinpoint the factors contributing to narrativization, and documents the relationship between the narrative response and other dimensions of musical experience captured by classical music psychology experiments. Some of this work closely examines the role of musical structures and patterns, directly manipulating the content of individual excerpts and examining the consequences for narrativization. It isolates cues in the music that convey topical associations through the accumulation of exposures to that pattern in particular contexts from structural cues, such as contrast that might impact narrativization independent of enculturation. Another future line of research targets the individuals who, despite the cross-cultural prevalence of NE with instrumental music, themselves showed little to no such engagement. This research attempts to better characterize these individuals and understand the ways that they differ from participants who sometimes or often reported narrative experiences.

The cross-cultural frequency of narrative response demonstrated by the data in this study, as well as the cross-cultural divergences in the excerpts that elicit this response, suggest that the narrative potential of music lies at the core of how apparently

abstract sound patterns can come to have rich meaning and significance in people's lives. This exposes a new alignment between music and language, and argues that narrative can emerge in multiple modalities, reinforcing its centrality in human strategies for understanding the world. Moreover, the rich cultural dependencies that shape which sounds come to be heard this way demonstrate that no reductionist account can link sound patterns and perceived meaning for real-world musical excerpts; rather, the affordances and limitations of the cognitive apparatus intertwine with experience and culture to shape musical meaning.

Data availability

To be available at <https://taplab.psy.msu.edu/>.

Received: 21 August 2019; Accepted: 7 November 2019;

Published online: 26 November 2019

Notes

1 We conducted an exploratory analysis that examined potential gender differences in narrative engagement for the Arkansas and Dimen samples. A comparison of NE scores for males and females revealed no significant gender differences in NE for either the Arkansas sample, $t(316) = 1.64$, $p = 0.1$, or the Dimen sample, $t(139) = -0.11$, $p = 0.91$.

References

- Abbott HP (2008) *The Cambridge introduction to narrative*. Cambridge University Press, Cambridge, UK
- Almén B (2008) *A theory of musical narrative*. Indiana University Press, Bloomington, IN
- Besson M, Friederici AD (1998) Language and music: a comparative view. *Music Percept* 16:1–9
- Brown S, Martinez MJ, Parsons LM (2006) Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *Eur J Neurosci* 23:2791–2803
- Dilley LC, McAuley JD (2008) Distal prosodic context affects word segmentation and lexical processing. *J Mem Lang* 59:294–311
- Fauconnier G, Turner MB (2008) *Rethinking metaphor*. In: Gibbs RW (ed) *Cambridge handbook of metaphor and thought*. Cambridge University Press, New York, NY, pp 53–66
- Forster EM (1910) *Howards end*. Edward Arnold and Company, London, UK
- Garrido S, Davidson J (2019) *Music, nostalgia and memory: historical and psychological perspectives*. Palgrave Macmillan Memory Studies, New York, NY
- Geertz C (1973) *The interpretation of cultures*. Basic Books, New York, NY
- Gourlay KA (1984) The non-universality of music and the universality of non-music. *World Music* 26:25–23
- Hannon EE, Trehub SE (2005) Tuning in to musical rhythms: infants learn more readily than adults *Proc Natl Acad Sci USA* 102:12639–12643
- Heider F, Simmel M (1944) An experimental study of apparent behavior. *Am J Psychol* 57:243–259
- Henrich J, Heine SJ, Norenzayan A (2010) The weirdest people in the world? *Behav Brain Sci* 33:61–83
- Herman D, Phelan J, Rabinowitz PJ, Richardson B, Warhol RR (2012) *Narrative theory: core concepts and critical debates*. Ohio State University Press, Columbus, OH
- Huovinen E, Kaila A-K (2015) The semantics of musical topoi: an empirical approach. *Music Percept* 33:217–243
- Hui MK, Dubé L, Chebat J-C (1997) The impact of music on consumers' reactions to waiting for services. *J Retail* 73:87–104
- Jackendoff R, Lerdahl F (2006) The capacity for music: what is it, and what's special about it? *Cognition* 100:33–72
- Juslin PN, Sloboda JA (eds) (2011) *Handbook of music and emotion: theory, research, applications*. Oxford University Press, New York, NY
- Juslin PN, Västfjäll D (2008) Emotional responses to music: the need to consider underlying mechanisms. *Behav Brain Sci* 31:559–575
- Koelsch S, Kasper E, Sammler D, Schulze K, Gunter T, Friederici AD (2004) Music, language and meaning: brain signatures of semantic processing. *Nat Neurosci* 7:302–307
- Margulis EH (2017) An exploratory study of narrative experiences of music. *Music Percept* 35:235–248
- McAdams DP (2006) *The redemptive self: stories Americans live by*. Oxford University Press, New York, NY

- Patel AD (2008) *Music, language, and the brain*. Oxford University Press, New York, NY
- Gordon RL, Jacobs MS, Schuele CM, McAuley JD (2015) Perspectives on the rhythm-grammar link and its implications for typical and atypical language development. *Ann New Y Acad Sci* 1337:16–25
- Ramsey SR (1989) *The languages of China*. Princeton University Press, Princeton, NJ
- Sloboda JA (2011) Music in everyday life: the role of emotions. In: Juslin PN, Sloboda JA (eds) *Handbook of music and emotion: theory, research, applications*. Oxford University Press, New York, NY, pp 493–514
- Rentfrow PJ, Gosling SD (2003) The do re mi's of everyday life: the structure and personality correlates of music preferences. *J Personal Soc Psychol* 84:1236–1256
- Tagg P, Clarida B (2003) *Ten little title tunes: towards a musicology of the mass media*. The Mass Media Musicologist's Press, New York, NY
- Vaux B, Golder S (2003) *The Harvard dialect survey*. Harvard University Linguistics Department, Cambridge, MA
- Vines BW, Krumhansl CL, Wanderley MM, Dalca IM, Levitin DJ (2011) Music to my eyes: cross-modal interactions in the perception of emotions in musical performance. *Cognition* 118:157–170
- Wieland EA, McAuley JD, Dilley LC, Chang SE (2015) Evidence for a rhythm perception deficit in children who stutter. *Brain Lang* 144:26–34
- Wong PC, Chan AH, Roy A, Margulis EH (2011) The bimusical brain is not two monomusical brains in one: evidence from musical affective processing. *J Cogn Neurosci* 23:4082–4093

Acknowledgements

Senyao Shen, Yuchen He, Xuepei Tang, Jieqiong Che, Shuya Ma, Mingwei Zhao, Joe Lau, and Nan Zhao helped with data collection and translation for Dimen participants. Aaron Judd helped select the Chinese musical excerpts. Catherine Ingram provided insight into Dong culture and music. Natalie M. Phillips contributed numerous discussions and many helpful insights about the study and data, and the members of the Timing, Attention and Perception Lab and the Digital Humanities and Literary Cognition Lab at Michigan State University contributed many helpful comments about various aspects of this work. Special thanks go to Mr. LEE Wai Kit and the staff at the Dimen Dong Eco-Museum for making data collection possible, and to the people in Dimen who participated in this research. This research was supported by the Division of Behavioral and Cognitive Sciences of the National Science Foundation, Award Numbers 1734025 (PI: EHM) and 1734063 (PI: JDM).

Author contributions

EHM, JDM, and PCMW conceptualized and supervised the project, developed the methodology, and acquired the funding. EHM, PC-MW, and JDM performed the investigation and shared the project administration, and JDM conducted the formal analysis, validation, and visualization. RS-G helped program the software. EHM and JDM wrote the paper with substantial review and editing contributions from PCMW.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to E.H.M. or J.D.M.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019